

IMPROVED MODULI OF CONTINUITY FOR DEGENERATE PHASE TRANSITIONS

UGO GIANAZZA, NAIAN LIAO, AND JOSÉ MIGUEL URBANO

ABSTRACT. We substantially improve in two scenarios the current state-of-the-art modulus of continuity for weak solutions to the N -dimensional, two-phase Stefan problem featuring a p -degenerate diffusion: for $p = N \geq 3$, we sharpen it to

$$\omega(r) \approx \exp(-c|\ln r|^{\frac{1}{N}});$$

for $p > \max\{2, N\}$, we derive an unexpected Hölder modulus.

1. INTRODUCTION

The classical two-phase Stefan problem is an archetypal free boundary problem that models a phase transition at constant temperature. It consists of solving the heat equation (or nonlinear variants of it) in the solid and liquid phases, coupled with the so-called Stefan condition at the *a priori* unknown interface separating them. This condition corresponds to an energy balance, prescribing the proportionality between the jump of the heat flux across the free boundary and its local velocity.

In this paper, we are interested in a phase transition problem involving a degenerate diffusion of p -Laplacian-type, for $p > 2$. In its weak form, for which any explicit reference to the free boundary is absent, such a problem can be formulated as

$$\partial_t \beta(u) - \operatorname{div} \mathbf{A}(x, t, u, Du) \ni 0 \quad \text{weakly in } E_T, \quad (1.1)$$

for a function $u : E_T \rightarrow \mathbb{R}$ representing the temperature. Here, $E_T \stackrel{\text{def}}{=} E \times (0, T]$, for some open set $E \subset \mathbb{R}^N$, $N \in \mathbb{N}$ and $T > 0$, whereas $\beta(\cdot)$ is the maximal monotone graph defined by

$$\beta(s) = \begin{cases} s & \text{if } s > 0, \\ [-\nu, 0] & \text{if } s = 0, \\ s - \nu & \text{if } s < 0, \end{cases} \quad (1.2)$$

for a positive constant ν , the latent heat of the phase transition, representing the aforementioned proportionality ratio. The function $\mathbf{A}(x, t, z, \xi) : E_T \times \mathbb{R} \times \mathbb{R}^N \rightarrow \mathbb{R}^N$ is a *Carathéodory* function, satisfying the structure conditions

$$\begin{cases} \mathbf{A}(x, t, z, \xi) \cdot \xi \geq C_0 |\xi|^p \\ |\mathbf{A}(x, t, z, \xi)| \leq C_1 |\xi|^{p-1} \end{cases} \quad \text{a.e. } (x, t) \in E_T, \forall z \in \mathbb{R}, \forall \xi \in \mathbb{R}^N, \quad (1.3)$$

Date: August 21, 2024.

2020 Mathematics Subject Classification. Primary 35K65, 35K92. Secondary 35B65, 76S05, 80A22.

Key words and phrases. Two-phase transition, parabolic p -Laplacian, intrinsic scaling, modulus of continuity, p -capacity.

where $0 < C_o \leq C_1$ are given positive constants.

Contrary to what was a common belief at the inception of the modern mathematical theory of the problem, the temperature u turns out to be a continuous function. This was proven independently by Caffarelli-Evans [4], DiBenedetto [6], Sacks [21], and Ziemer [25] in the early 1980s, much to the surprise of the Russian school that developed the weak formulation and the corresponding existence theory in the 1960s (see [20], and also [24]). The result would eventually be extended in [22] to the case $p > 2$ treated here.

The quantification of these qualitative results, *i.e.*, the search for a modulus of continuity for the temperature u , has a history of its own. Besides its intrinsic interest, the knowledge of a quantitative modulus of continuity is important for applications. Indeed, if we define the (weak) free boundary as $\partial[u > 0]$, there is a direct connection between the regularity of this interface and the modulus of continuity of u , as discussed, for example, in [5, Section 6].

A modulus of the type

$$\omega(r) \approx \left[\ln \left| \ln \left(\frac{r}{R} \right) \right| \right]^{-\sigma}, \quad r \in (0, R),$$

for some $\sigma > 0$, which follows from the reasoning in [6, 4], would eventually appear explicitly in [9, Remark 3.1] and [7]. It remained the state-of-the-art for 30 years until the next decisive result in [1] improved the modulus to

$$\omega(r) \approx \left| \ln \left(\frac{r}{R} \right) \right|^{-\sigma(N,p)}, \quad r \in (0, R),$$

including the degenerate case $p > 2$, discarding an iteration of the logarithm and determining the precise value of the exponent σ in terms of the data of the problem: in the case $2 < p < N$, for example, $\sigma = \frac{p}{N+p}$. The conjecture stated in [1] that the sharp, optimal modulus of continuity for the two-phase Stefan problem had been obtained turned out to be speculative, and we disprove it here, obtaining a better modulus in two scenarios: for $p = N$ in dimensions $N = 3, 4, \dots$, the modulus is upgraded to

$$\omega(r) \approx \exp \left(-c \left| \ln \left(\frac{r}{R} \right) \right|^{\frac{1}{N}} \right), \quad r \in (0, R);$$

for dimensions $N = 1, 2$, and for $p > N$ in dimensions $N = 3, 4, \dots$, the modulus is improved to a Hölder modulus of type

$$\omega(r) \approx \left(\frac{r}{R} \right)^\gamma, \quad r \in (0, R).$$

These findings constitute the object of our main result.

Theorem 1.1. *Let u be a bounded, local weak solution to the Stefan problem (1.1)–(1.3) and let*

$$\mathbb{M} = \max \left\{ 1, \operatorname{osc}_{E_T} u \right\}.$$

Consider a compact subset \mathcal{K} of E_T and denote by R the p -parabolic distance of \mathcal{K} to the parabolic boundary of E_T , cf. (2.1).

(I) When $p = N \geq 3$, there exist constants $c, \sigma \in (0, 1)$, depending on $\{\nu, N, C_o, C_1\}$, such that for every $(x_i, t_i) \in \mathcal{K}$, $i = 0, 1$, satisfying

$$|x_o - x_1| + \mathbb{M}^{\frac{N-2}{N}} |t_o - t_1|^{\frac{1}{N}} \leq \sigma R,$$

we have

$$|u(x_o, t_o) - u(x_1, t_1)| \leq \mathbb{M} \exp \left(-c \left| \ln \left(\frac{|x_o - x_1| + \mathbb{M}^{\frac{N-2}{N}} |t_o - t_1|^{\frac{1}{N}}}{\sigma R} \right) \right|^{\frac{1}{N}} \right).$$

(II) When $p > \max\{2, N\}$, there exist constants $\gamma \in (0, 1)$ and $C > 1$, depending on $\{N, p, C_o, C_1\}$ but independent of ν , such that for every $(x_i, t_i) \in \mathcal{K}$, $i = 0, 1$, satisfying

$$|x_o - x_1| + \mathbb{M}^{\frac{p-2}{p}} |t_o - t_1|^{\frac{1}{p}} \leq R,$$

we have

$$|u(x_o, t_o) - u(x_1, t_1)| \leq C \mathbb{M} \left(\frac{|x_o - x_1| + \mathbb{M}^{\frac{p-2}{p}} |t_o - t_1|^{\frac{1}{p}}}{R} \right)^\gamma.$$

For the remaining case, namely $2 < p < N$, $N \geq 3$, our approach yields a logarithmic type modulus of continuity like the one obtained in [1]. Therefore, the significance of our contribution lies in bringing previously hidden structural properties of solutions to light and providing a unified perspective on the issue in all scenarios. Table 1 summarises the current state-of-the-art in terms of the modulus of continuity (MoC) for the two-phase Stefan problem involving a degenerate diffusion of p -Laplacian-type, for $p > 2$.

p & N	MoC
$p > \max\{2, N\}$	$\omega^{(1)}(r) \approx r^\gamma$
$p = N \geq 3$	$\omega^{(2)}(r) \approx \exp(-c \ln r ^{\frac{1}{N}})$
$2 < p < N, N \geq 3$	$\omega^{(3)}(r) \approx \ln r ^{-\sigma}$

TABLE 1. Moduli of Continuity

Among these three types of moduli, the Hölder modulus of continuity $\omega^{(1)}(\cdot)$ is perhaps the most surprising one. Loosely speaking, it shows the p -Laplacian-type diffusion dictates the local behaviour of solutions as though the singularity of $\beta(\cdot)$ plays no role. The Hölder exponent γ determined by our method is implicit (see [19] for a slightly different approach), and it would be interesting to search for an explicit γ , even in the archetypal case of the p -Laplacian. On the other hand, the modulus $\omega^{(2)}(\cdot)$ represents a sort of borderline case; it was observed by Caffarelli and Friedman in their study [5] of the so-called “one-phase” Stefan problem and, recently, established in [18], in the case $p = N = 2$, under a general framework similar to that of this paper.

1.1. New techniques. The singularity of β occurs at $u = 0$; henceforth, it is natural to expect that the set $[u = 0]$ and its measure play a fundamental role in the derivation of the modulus of continuity. As is frequently the case when studying the regularity of solutions to the two-phase Stefan problem, the need to deal with such a measure leads to considering an alternative. However, a significant difference appears here: instead of taking into account the measure of $[u = 0]$ inside a proper cylinder, we deal with measures inside time slices, that is, sets of the form $[u(\cdot, t) = 0] \cap K_\varrho$, for t in a suitable interval I , where K_ϱ denotes the cube of wedge 2ϱ . Therefore, the alternative is stated as follows: either there exists a $\bar{t} \in I$ such that $|[u(\cdot, \bar{t}) = 0]|$ is small in K_ϱ , or for every $t \in I$, the set where $u(\cdot, t)$ is close to zero occupies a sizable portion of the cube.

The former possibility is the favourable one, and it is dealt with in Sections 4.1 and 4.3; roughly speaking, the smallness of the singularity set yields that the solution behaves as if it were the solution of a p -Laplacian-type equation, and its continuity can be obtained in a fairly standard way.

The latter occurrence is more complicated, and it is tackled in Sections 4.2 and 4.4 thanks to Proposition 3.1, which is a new *expansion of positivity* for non-negative super-solutions to parabolic p -Laplacian-type equations as (3.1), satisfying the structure conditions (1.3). It is precisely this result that allows us to discriminate the behaviour of solutions in terms of the dimension N . We believe it might be of independent interest.

To better understand the novelty, let us first consider the by-now classical expansion of positivity for such a super-solution v to parabolic p -Laplacian-type equations when $p > 2$ (see [10, Chapter 5, Proposition 7.1]): if, for some $k > 0$ and $\alpha \in (0, 1)$,

$$|[v(\cdot, 0) \geq k] \cap K_1| \geq \alpha |K_1|,$$

there exist constants $\xi \in (0, 1)$, and $b, d > 1$, depending on $\{N, p, C_o, C_1\}$, such that

$$v(\cdot, t) \geq \xi \alpha^d k \quad \text{in } K_2,$$

for all times

$$\frac{b}{(\xi \alpha^d k)^{p-2}} \leq t \leq \frac{2b}{(\xi \alpha^d k)^{p-2}}.$$

The crucial issue is the power-like dependence of the shrinking parameter $\xi \alpha^d$ on the measure-theoretical quantity α . In contrast, Proposition 3.1 yields that, given a super-solution v , if

$$|[v(\cdot, t) \geq k] \cap K_1| \geq \alpha |K_1|,$$

for all $t \in (0, (\bar{\delta} \text{cap}_p(K_{\alpha^2}, K_3) k)^{2-p}]$, for a proper $\bar{\delta}$ which depends only on the data $\{N, p, C_o, C_1\}$, there exist constants $\xi \in (0, 1)$ and $b > 1$, with the same dependence as $\bar{\delta}$, such that

$$v(\cdot, t) \geq \xi \text{cap}_p(K_{\alpha^2}, K_3) k \quad \text{in } K_2,$$

for all times

$$\frac{b}{(\xi \text{cap}_p(K_{\alpha^2}, K_3) k)^{p-2}} \leq t \leq \frac{2b}{(\xi \text{cap}_p(K_{\alpha^2}, K_3) k)^{p-2}}.$$

The different sets where the initial measure-theoretical lower bound is achieved are just technical differences and do not play any substantial role; the main point is that the shrinking parameter changes from $\xi \alpha^d$ to $\xi \text{cap}_p(K_{\alpha^2}, K_3)$ and, relying on the properties

of the p -capacity, we can dramatically improve the dependence with respect to α when $p \geq N$, and consequently obtain the sharpened moduli $\omega^{(1)}$ and $\omega^{(2)}$ of Table 1. The improvement of the shrinking parameter is based on the deep measure-theoretical Lemma 2.1, and a proper adaptation of the capacity estimates originally employed in [12]. A result similar to Proposition 3.1 was first used in [18] for non-degenerate parabolic equations; extending such a statement to the degenerate context $p > 2$ is far from trivial. Indeed, it requires careful use of the sophisticated intrinsic scaling method, typical of parabolic p -Laplacian-type equations, for which we refer to [8, 10, 23].

Finally, it is also worth mentioning that our approach forgoes the use of logarithmic estimates, unlike most previous works on the topic. The mechanism we employ to spread information in time is based on Lemma 2.3, which follows from energy estimates alone.

1.2. Organization of the paper. We organize the paper as follows. In Section 2, we introduce the basic notation and definitions and collect some auxiliary results. In Section 3, we present the new expansion of positivity and its proof, whereas the proof of the main theorem is given in Section 4.

2. PRELIMINARY RESULTS

This section gathers some useful pieces of notation, the main definitions, and several auxiliary results that will be instrumental in the main proofs. They comprise a measure-theoretical lemma, energy estimates and three De Giorgi-type lemmata; for the third, a detailed proof is included since it generalizes to any $p > 1$ the corresponding result for $p = 2$.

2.1. Notation and definitions. For $\varrho > 0$ and $y \in \mathbb{R}^N$, denote by $K_\varrho(y)$ the cube of edge 2ϱ , centred at y and with faces parallel to the coordinate planes, and by $B_\varrho(y)$ the ball of radius ϱ centred at y . If y is the origin, let $K_\varrho(0) = K_\varrho$, and $B_\varrho(0) = B_\varrho$. Most of the statements are given in terms of cubes, but we briefly refer to balls in Appendix A when considering simple, explicit examples of p -capacity.

The parabolic boundary of the cylinder E_T is given by the union of its lateral and initial boundary

$$\partial_{\text{par}} E_T \stackrel{\text{def}}{=} (\overline{E} \times \{0\}) \cup (\partial E \times (0, T]).$$

For any compact subset \mathcal{K} of E_T , the p -parabolic distance of \mathcal{K} to the parabolic boundary $\partial_{\text{par}} E_T$ is defined by

$$p - \text{dist}_{\text{par}}(\mathcal{K}, \partial_{\text{par}} E_T) \stackrel{\text{def}}{=} \inf_{\substack{(x,t) \in \mathcal{K} \\ (y,s) \in \partial_{\text{par}} E_T}} \left(\max \left\{ \max_{1 \leq i \leq N} |x_i - y_i|, \mathbb{M}^{\frac{p-2}{p}} |t - s|^{\frac{1}{p}} \right\} \right). \quad (2.1)$$

As is usual with structure conditions like those in (1.3), the definition of p -parabolic distance depends on the solution itself. It is not hard to see that the previous definition is equivalent to the one usually employed with the parabolic p -Laplacian, namely

$$p - \text{dist}_{\text{par}}(\mathcal{K}, \partial_{\text{par}} E_T) = \inf_{\substack{(x,t) \in \mathcal{K} \\ (y,s) \in \partial_{\text{par}} E_T}} \left(|x - y| + \mathbb{M}^{\frac{p-2}{p}} |t - s|^{\frac{1}{p}} \right).$$

Using (2.1) simplifies some of the computations at the end of this work (cf. Subsection 4.6).

For $R, S, \varrho, \theta > 0$, the backward and forward cylinders with vertex (x_o, t_o) in \mathbb{R}^{N+1} are defined by

$$\begin{cases} (x_o, t_o) + Q_{R,S} \stackrel{\text{def}}{=} K_R(x_o) \times (t_o - S, t_o], \\ (x_o, t_o) + Q_\varrho(\theta) \stackrel{\text{def}}{=} K_\varrho(x_o) \times (t_o - \theta\varrho^p, t_o], \\ (x_o, t_o) + Q_\varrho^+(\theta) \stackrel{\text{def}}{=} K_\varrho(x_o) \times (t_o, t_o + \theta\varrho^p]. \end{cases}$$

Functions truncated by $k \in \mathbb{R}$ are defined as

$$(u - k)_+ \stackrel{\text{def}}{=} \max\{u - k, 0\}, \quad (u - k)_- \stackrel{\text{def}}{=} \max\{k - u, 0\}.$$

For a Lebesgue measurable set $A \subset \mathbb{R}^N$, we denote its measure by $|A|$. In various estimates, we use γ to represent a generic positive constant.

We now introduce the precise notions of local weak sub-solution, local weak super-solution, and local weak solution, which we will deal with throughout the paper. The relationship between classical and weak formulations of the two-phase Stefan problem is an interesting issue in itself, for which we refer to [24].

Definition 2.1. *A function*

$$u \in L_{\text{loc}}^\infty(0, T; L_{\text{loc}}^2(E)) \cap L_{\text{loc}}^p(0, T; W_{\text{loc}}^{1,p}(E))$$

is a local weak sub(super)-solution to the Stefan problem (1.1) with the structure conditions (1.3), if, for every compact set $\mathcal{K} \subset E$ and every sub-interval $[t_1, t_2] \subset (0, T]$, there is a selection $v \subset \beta(u)$, i.e.,

$$\{(z, v(z)) : z \in E_T\} \subset \{(z, \beta[u(z)]) : z \in E_T\},$$

such that

$$\int_{\mathcal{K}} v \zeta \, dx \Big|_{t_1}^{t_2} + \iint_{\mathcal{K} \times (t_1, t_2)} [-v \partial_t \zeta + \mathbf{A}(x, t, u, Du) \cdot D\zeta] \, dx dt \leq (\geq) 0,$$

for all non-negative testing functions

$$\zeta \in W_{\text{loc}}^{1,2}(0, T; L^2(\mathcal{K})) \cap L_{\text{loc}}^p(0, T; W_o^{1,p}(\mathcal{K})).$$

A function that is both a local weak sub-solution and a local weak super-solution is termed a local weak solution.

The above notion of solution, though natural from the viewpoint of existence theory, renders an issue when one attempts to use the solution as a testing function since the time derivative does not exist in the Sobolev sense. To confront it, we will consider the regularized version of the Stefan problem (1.1):

$$\partial_t \beta_\epsilon(u) - \operatorname{div} \mathbf{A}(x, t, u, Du) = 0, \quad \text{weakly in } E_T. \quad (2.2)$$

Here, we define the mollification of β by

$$\beta_\epsilon(s) \stackrel{\text{def}}{=} s + \nu H_\epsilon(s),$$

for a parameter $\epsilon \in (0, 1)$ and

$$H_\epsilon(s) \stackrel{\text{def}}{=} \begin{cases} 0, & s > 0, \\ \frac{1}{\epsilon}s, & -\epsilon \leq s \leq 0, \\ -1, & s < -\epsilon. \end{cases}$$

The notion of solution for (2.2) can be found in [8, Chapter II]. We will proceed with the assumption that local solutions to the Stefan problem (1.1) can be approximated by a sequence of solutions to (2.2) locally uniformly. This is a standard assumption, cf. [11, 1, 13].

It is well-known that, for any fixed ϵ , solutions to (2.2) are locally Hölder continuous; see [8, Chapter III, IV]. However, such Hölder estimates will not be retained as $\epsilon \rightarrow 0$. We aim to identify estimates on the modulus of continuity that are stable as $\epsilon \rightarrow 0$. In this way, the Stefan problem (1.1), as a limiting case of the regularized problem (2.2), enjoys the same kind of modulus of continuity. This kind of regularizing scheme also appears in the study of gradient regularity for congested traffic dynamics [2, 3].

2.2. A measure-theoretical lemma. The following result exhibits the clustering of positivity for Sobolev functions in a measure-theoretical sense.

Lemma 2.1. *Let $u \in W^{1,1}(K_\rho)$ satisfy*

$$\int_{K_\rho} |Du| \, dx \leq \gamma \frac{k}{\rho} \quad \text{and} \quad |[u > k] \cap K_\rho| \geq \alpha |K_\rho|,$$

for some positive k, γ , and $\alpha \in (0, 1)$. Then, for every $\delta, \lambda \in (0, 1)$, there exist $\epsilon = \epsilon(\alpha, \delta, \gamma, \lambda, N) \in (0, 1)$ and $K_{\epsilon\rho}(y) \subset K_\rho$, such that

$$|[u > (1 - \lambda)k] \cap K_{\epsilon\rho}(y)| > (1 - \delta)|K_{\epsilon\rho}|.$$

Remark 2.1. Following the various steps of the proof (for which we refer to [10, Chapter 2, Lemma 3.1]), the dependence of the reducing parameter ϵ on the measure-theoretical parameter α , and on the constant γ appearing in the assumptions of the lemma can be traced to be of the form

$$\epsilon = \bar{c} \alpha^2 \tag{2.3}$$

for a constant $\bar{c} = \frac{\lambda\delta}{\gamma C}$, where $C > 1$ depends only on N , and is independent of α .

2.3. Energy estimates. Consider a reference cylinder $(y, \tau) + Q_{R,S} \subset E_T$. For simplicity, we omit the reference point in the sequel and denote the cylinder as $Q_{R,S}$. The following result has been proven in [13, Proposition 2.1].

Proposition 2.1. *Let $p > 1$ and consider a local weak sub(super)-solution u to the Stefan problem (2.2) with structure conditions (1.3) in E_T . There exists a positive constant $\gamma = \gamma(p, C_0, C_1)$, such that for all cylinders $Q_{R,S} \subset E_T$, every $k \in \mathbb{R}$, and every non-negative, piecewise smooth cutoff function ζ vanishing on $\partial K_R(y) \times (\tau - S, \tau]$, there holds*

$$\text{ess sup}_{\tau-S < t < \tau} \frac{1}{2} \int_{K_R(y) \times \{t\}} \zeta^p (u - k)_\pm^2 \, dx$$

$$\begin{aligned}
& + \nu \operatorname{ess\,sup}_{\tau-S < t < \tau} \int_{K_R(y) \times \{t\}} \left(\int_0^{(u-k)_\pm} H'_\epsilon(k \pm s) s \, ds \right) \zeta^p \, dx \\
& + \iint_{Q_{R,S}} \zeta^p |D(u-k)_\pm|^p \, dx dt \\
\leq & \gamma \iint_{Q_{R,S}} [(u-k)_\pm^p |D\zeta|^p + (u-k)_\pm^2 |\partial_t \zeta^p|] \, dx dt \tag{2.4} \\
& + \nu \iint_{Q_{R,S}} \left(\int_0^{(u-k)_\pm} H'_\epsilon(k \pm s) s \, ds \right) |\partial_t \zeta^p| \, dx dt \\
& + \frac{1}{2} \int_{K_R(y) \times \{\tau-S\}} \zeta^p (u-k)_\pm^2 \, dx \\
& + \nu \int_{K_R(y) \times \{\tau-S\}} \left(\int_0^{(u-k)_\pm} H'_\epsilon(k \pm s) s \, ds \right) \zeta^p \, dx.
\end{aligned}$$

Here, we adopted the convention of interpreting \pm as $+$ for sub-solutions and as $-$ for super-solutions. This convention also applies to the statements to follow.

2.4. De Giorgi-type lemmata. For $Q_{R,S} \subset E_T$, introduce the numbers μ^\pm and ω satisfying

$$\mu^+ \geq \operatorname{ess\,sup}_{Q_{R,S}} u, \quad \mu^- \leq \operatorname{ess\,inf}_{Q_{R,S}} u, \quad \omega \geq \mu^+ - \mu^-.$$

We discuss three lemmata resulting from the energy estimate (2.4) and De Giorgi-type iterations. It is worth mentioning that they all hold for $p > 1$.

The first De Giorgi-type lemma is the same as [13, Lemma 2.1]; see [17, Lemma 2.1] for a proof.

Lemma 2.2. *Let u be a local weak sub(super)-solution to the Stefan problem (2.2) with structure conditions (1.3) in E_T . Let $\delta, \xi \in (0, 1)$, and set $\theta = \delta(\xi\omega)^{2-p}$. There exists a constant $c_o \in (0, 1)$, depending only on the data $\{\nu, N, p, C_o, C_1\}$, such that if $\xi\omega \leq 1$ and if*

$$|[\pm(\mu^\pm - u) \leq \xi\omega] \cap [(x_o, t_o) + Q_\rho(\theta)]| \leq c_o \delta^{\frac{N}{p}} (\xi\omega)^{\frac{N+p}{p}} |Q_\rho(\theta)|,$$

then

$$\pm(\mu^\pm - u) \geq \frac{1}{2}\xi\omega, \quad \text{a.e. in } (x_o, t_o) + Q_{\frac{1}{2}\rho}(\theta),$$

provided the cylinder $(x_o, t_o) + Q_\rho(\theta)$ is included in $Q_{R,S}$.

Remark 2.2. Tracing the various constants in the proof gives that $c_o = \tilde{c}_o \nu_*^{-(N+p)/p}$, where $\nu_* = \max\{1, \nu\}$ and $\tilde{c}_o \in (0, 1)$ depends only on $\{N, p, C_o, C_1\}$.

The second De Giorgi-type lemma can be retrieved from [13, Lemma 2.2] or [19, Lemma 2.1].

Lemma 2.3. *Let u be a local weak sub(super)-solution to the Stefan problem (2.2) with structure conditions (1.3) in E_T . Assume that, for some $\xi \in (0, 1)$, there holds*

$$\pm(\mu^\pm - u(\cdot, t_1)) \geq \xi\omega, \quad \text{a.e. in } K_\rho(x_o).$$

There exists a constant $\gamma_o \in (0, 1)$, depending only on the data $\{N, p, C_o, C_1\}$ and independent of ν , such that, for any $\theta > 0$, if

$$|[\pm(\mu^\pm - u) \leq \xi\omega] \cap [(x_o, t_1) + Q_\varrho^+(\theta)]| \leq \frac{\gamma_o(\xi\omega)^{2-p}}{\theta} |Q_\varrho^+(\theta)|,$$

then

$$\pm(\mu^\pm - u) \geq \frac{1}{2}\xi\omega, \quad \text{a.e. in } K_{\frac{1}{2}\varrho}(x_o) \times (t_1, t_1 + \theta\varrho^p),$$

provided the cylinder $(x_o, t_1) + Q_\varrho^+(\theta)$ is included in $Q_{R,S}$.

The third De Giorgi-type lemma generalizes to any $p > 1$ the corresponding result for $p = 2$ obtained in [18, Lemma 2.2].

Lemma 2.4. *Let u be a local weak sub(super)-solution to the Stefan problem (2.2) with structure conditions (1.3) in E_T . For $\xi \in (0, 1)$, set $\theta = (\xi\omega)^{2-p}$. There exist $c_1, \delta \in (0, 1)$, depending only on the data $\{\nu, N, p, C_o, C_1\}$, such that if $\xi\omega \leq 1$ and if*

$$|[\pm(\mu^\pm - u(\cdot, t_1)) \leq \xi\omega] \cap K_{2\varrho}(x_o)| \leq c_1(\xi\omega)^{\frac{N+2p}{p}} |K_{2\varrho}|, \quad (2.5)$$

then

$$\pm(\mu^\pm - u) \geq \frac{1}{4}\xi\omega, \quad \text{a.e. in } K_{\frac{1}{2}\varrho}(x_o) \times (t_1 + (1 - 2^{-p})\delta\theta\varrho^p, t_1 + \delta\theta\varrho^p), \quad (2.6)$$

provided $K_{2\varrho}(x_o) \times (t_1, t_1 + \delta\theta\varrho^p)$ is included in $Q_{R,S}$.

Proof. Let us deal with the case of super-solutions only since the other case is similar. Without loss of generality, we assume $(x_o, t_1) = (0, 0)$, as we can always resort to this case by translation.

We start by defining the sequences, indexed over the set of non-negative integers \mathbb{N}_0 ,

$$k_n = \mu^- + \frac{\xi\omega}{2} + \frac{\xi\omega}{2^{n+1}},$$

and

$$\varrho_n = \varrho + \frac{\varrho}{2^n},$$

and the corresponding cubes and cylinders

$$K_n = K_{\varrho_n} \quad \text{and} \quad Q_n = K_n \times (0, \delta\theta\varrho^p],$$

where $\delta > 0$ will be determined in due course. We also consider the sequence of middle points

$$\tilde{\varrho}_n = \frac{\varrho_n + \varrho_{n+1}}{2},$$

with $\tilde{K}_n = K_{\tilde{\varrho}_n}$ and $\tilde{Q}_n = \tilde{K}_n \times (0, \delta\theta\varrho^p]$.

We now write the energy estimate (2.4) over Q_n , for a cutoff function $\zeta \in C_o^\infty(K_n)$ such that $\zeta(x) \in [0, 1]$, $\zeta \equiv 1$ in \tilde{K}_n and $|D\zeta| \leq 2^{n+4}/\varrho$, obtaining

$$\begin{aligned} & \operatorname{ess\,sup}_{0 < t < \delta\theta\varrho^p} \int_{\tilde{K}_n \times \{t\}} (u - k_n)_-^2 \, dx + \iint_{\tilde{Q}_n} |D(u - k_n)_-|^p \, dxdt \\ & \leq \gamma \frac{2^{np}}{\varrho^p} \iint_{Q_n} (u - k_n)_-^p \, dxdt + \int_{K_n \times \{0\}} (u - k_n)_-^2 \, dx \\ & \quad + 2\nu \int_{K_n \times \{0\}} (u - k_n)_- \, dx \end{aligned} \quad (2.7)$$

$$\begin{aligned}
&\leq \gamma \frac{2^{np}}{\varrho^p} (\xi\omega)^p |[u < k_n] \cap Q_n| + (\xi\omega)^2 |[u(\cdot, 0) < k_n] \cap K_n| \\
&\quad + 2\nu\xi\omega |[u(\cdot, 0) < k_n] \cap K_n| \\
&\leq \gamma \frac{2^{np}}{\varrho^p} (\xi\omega)^p |[u < k_n] \cap Q_n| + 2\nu_*\xi\omega |[u(\cdot, 0) < k_n] \cap K_n| \\
&\leq \gamma \frac{2^{np}}{\varrho^p} (\xi\omega)^p |A_n| + 2^{N+1}\nu_*c_1(\xi\omega)^{\frac{N+3p}{p}} |K_n|,
\end{aligned}$$

with $\nu_* \stackrel{\text{def}}{=} \max\{1, \nu\}$.

We used the fact that $\xi\omega < 1$ and thus $(\xi\omega)^2 < \xi\omega$ and, in the last estimate, the following consequence of (2.5), due to $\varrho < \varrho_n \leq 2\varrho$:

$$\begin{aligned}
|[u(\cdot, 0) \leq k_n] \cap K_n| &\leq |[u(\cdot, 0) \leq \mu^- + \xi\omega] \cap K_{2\varrho}| \\
&\leq c_1(\xi\omega)^{\frac{N+2p}{p}} |K_{2\varrho}| \\
&\leq c_1(\xi\omega)^{\frac{N+2p}{p}} 2^N |K_n|,
\end{aligned}$$

for any $n \in \mathbb{N}_0$. We also denoted

$$A_n \stackrel{\text{def}}{=} [u < k_n] \cap Q_n.$$

We now consider an alternative: either, for some $n \in \mathbb{N}_0$,

$$|A_n| \leq \frac{c_1(\xi\omega)^{\frac{N+p}{p}}}{\delta} |Q_n| \quad (2.8)$$

or we have

$$|A_n| > \frac{c_1(\xi\omega)^{\frac{N+p}{p}}}{\delta} |Q_n|, \quad (2.9)$$

for every $n \in \mathbb{N}_0$. If (2.8) holds, assume δ is already chosen in terms of the data $\{\nu, N, p, C_o, C_1\}$, and choose c_1 as

$$c_1 = c_o 2^{-N - \frac{N}{p} - 1} \delta^{\frac{N}{p} + 1}, \quad (2.10)$$

where c_o is determined in Lemma 2.2. Noticing again that $\varrho < \varrho_n \leq 2\varrho$ and that $k_n > \mu^- + \frac{1}{2}\xi\omega$, we readily get

$$|[u < \mu^- + \frac{1}{2}\xi\omega] \cap [(0, \delta\theta\varrho^p) + Q_\varrho(\delta\theta)]| \leq |A_n| \leq c_o \delta^{\frac{N}{p}} (\frac{1}{2}\xi\omega)^{\frac{N+p}{p}} |Q_\varrho(\delta\theta)|.$$

Now, Lemma 2.2 gives

$$u \geq \mu^- + \frac{1}{4}\xi\omega \quad \text{in } K_{\frac{1}{2}\varrho} \times ((1 - 2^{-p})\delta\theta\varrho^p, \delta\theta\varrho^p]. \quad (2.11)$$

If, alternatively, (2.9) holds, we conclude from the energy estimate (2.7) that, for every $n \in \mathbb{N}_0$,

$$\begin{aligned}
&\text{ess sup}_{0 < t < \delta\theta\varrho^p} \int_{\tilde{K}_n \times \{t\}} (u - k_n)_-^2 dx + \iint_{\tilde{Q}_n} |D(u - k_n)_-|^p dx dt \\
&\leq (\gamma 2^{np} + 2^{N+1}\nu_*) \frac{(\xi\omega)^p}{\varrho^p} |A_n|
\end{aligned} \quad (2.12)$$

$$\leq \gamma 2^{np} \nu_* \frac{(\xi\omega)^p}{\varrho^p} |A_n|,$$

for some $\gamma = \gamma(N, C_o, C_1) > 1$, since (2.9) implies

$$c_1(\xi\omega)^{\frac{N+3p}{p}} |K_n| < \frac{(\xi\omega)^p}{\varrho^p} |A_n|.$$

We now run De Giorgi's iteration scheme. Take a cutoff function $\phi \in C_o^\infty(\tilde{K}_n)$ such that $\phi(x) \in [0, 1]$, $\phi \equiv 1$ in K_{n+1} and $|D\phi| \leq 2^{n+4}/\varrho$, to get, using Hölder's inequality and the parabolic Sobolev inequality (see [8, Chapter I, Proposition 3.1]),

$$\begin{aligned} \frac{\xi\omega}{2^{n+2}} |A_{n+1}| &= (k_n - k_{n+1}) |A_{n+1}| \\ &\leq \iint_{Q_{n+1}} (u - k_n)_- \, dxdt \\ &\leq \iint_{\tilde{Q}_n} \phi(u - k_n)_- \, dxdt \\ &\leq \left[\iint_{\tilde{Q}_n} |\phi(u - k_n)_-|^{\frac{p(N+2)}{N}} \, dxdt \right]^{\frac{N}{p(N+2)}} |A_n|^{1 - \frac{N}{p(N+2)}} \\ &\leq \gamma \left(\iint_{\tilde{Q}_n} |D[\phi(u - k_n)_-]|^p \, dxdt \right)^{\frac{N}{p(N+2)}} \\ &\quad \cdot \left(\operatorname{ess\,sup}_{0 < t < \delta\theta\varrho^p} \int_{\tilde{K}_n \times \{t\}} (u - k_n)_-^2 \, dx \right)^{\frac{1}{N+2}} |A_n|^{1 - \frac{N}{p(N+2)}} \\ &\leq \gamma \left(2^{np} \nu_* \frac{(\xi\omega)^p}{\varrho^p} |A_n| \right)^{\frac{N+p}{p(N+2)}} |A_n|^{1 - \frac{N}{p(N+2)}} \\ &= \gamma \left(2^{np} \nu_* \frac{(\xi\omega)^p}{\varrho^p} \right)^{\frac{N+p}{p(N+2)}} |A_n|^{1 + \frac{1}{N+2}}, \end{aligned}$$

for some $\gamma = \gamma(N, p, C_o, C_1) > 1$, where we also used estimate (2.12). We rewrite this inequality for

$$Y_n \stackrel{\text{def}}{=} \frac{|A_n|}{|Q_n|},$$

obtaining the recursive relation

$$Y_{n+1} \leq \gamma 2^{Npn} \nu_*^{\frac{N+p}{p(N+2)}} \delta^{\frac{1}{N+2}} Y_n^{1 + \frac{1}{N+2}}.$$

Using the fast geometric convergence of sequences (see [8, Chapter I, Lemma 4.1]), we conclude that $Y_n \rightarrow 0$ as $n \rightarrow \infty$ provided

$$Y_o \leq \frac{1}{2^{Np(N+2)^2} \gamma^{N+2} \nu_*^{\frac{N+p}{p}} \delta}.$$

This is granted by taking

$$\delta = \frac{1}{2^{Np(N+2)^2} \gamma^{N+2} \nu_*^{\frac{N+p}{p}}} \quad (2.13)$$

and we obtain

$$u \geq \mu^- + \frac{1}{2} \xi \omega \quad \text{in } K_\varrho \times (0, \delta \theta \varrho^p]. \quad (2.14)$$

The result now follows, since no matter which of the alternatives (2.8) or (2.9) holds, (2.6) can be obtained from either (2.11) or (2.14).

We conclude by tracing the constant dependence: with the choice of δ in (2.13), we obtain from (2.10), also taking into account the value of c_o in Remark 2.2,

$$c_1 = \frac{\tilde{c}_o(N, p, C_o, C_1)}{2^{N+1+N(N+2)^2(N+p)+\frac{N}{p}} \gamma^{(N+2)(1+\frac{N}{p})} \nu_*^{(1+\frac{N}{p})^2+(1+\frac{N}{p})}}.$$

This finishes the proof. \square

3. SUB/SUPER-SOLUTIONS TO THE PARABOLIC p -LAPLACIAN

In this section, we shall consider quasi-linear, parabolic partial differential equations of p -Laplacian-type

$$\partial_t u - \operatorname{div} \tilde{\mathbf{A}}(x, t, u, Du) = 0 \quad \text{weakly in } E_T, \quad (3.1)$$

where the function $\tilde{\mathbf{A}} : E_T \times \mathbb{R} \times \mathbb{R}^N \rightarrow \mathbb{R}^N$ satisfies (1.3) for $p > 2$. The corresponding notion of sub/super-solution is standard, and we refer to [8, Chapter II].

Before introducing the main result of this section, we state a helpful lemma that connects the Stefan problem with the parabolic p -Laplace equation; its proof can be retrieved from [1, Lemma 2.3] or [17, Lemma 4.3].

Lemma 3.1. *Let u be a local weak solution to the Stefan problem (2.2) with structure conditions (1.3) in E_T . Then, for any $k > 0$, the truncation $(u - k)_+$ is a local weak sub-solution to a parabolic equation of type (3.1) in E_T . Likewise, for any $k < -\epsilon$, the truncation $(u - k)_-$ is a local weak sub-solution to a parabolic equation of type (3.1) in E_T .*

The main result of this section is the following expansion of positivity. It translates measure-theoretical estimates to pointwise p -capacity estimates. The notion of p -capacity is discussed in Appendix A.

Proposition 3.1. *Let $k > 0$ and $\alpha \in (0, 1)$. Consider a non-negative, local weak super-solution v to the parabolic equation (3.1), with structure conditions (1.3) in E_T . There exist constants $\bar{c}, \bar{\delta}, \xi \in (0, 1)$ and $b > 1$, depending only on the data $\{N, p, C_o, C_1\}$, such that whenever*

$$|[v(\cdot, t) \geq k] \cap K_\varrho(x_o)| \geq \alpha |K_\varrho|$$

holds true for any $t \in (t_o, t_o + \bar{\theta} \varrho^p]$, where

$$\bar{\theta} \stackrel{\text{def}}{=} [k \bar{\delta} \operatorname{cap}_p(K_{\bar{c}\alpha^2}, K_3)]^{2-p},$$

then, we have

$$v \geq k \xi \operatorname{cap}_p(K_{\bar{c}\alpha^2}, K_3), \quad \text{a.e. in } K_\varrho(x_o) \times (t_o + b \xi^{2-p} \bar{\theta} \varrho^p, t_o + 2b \xi^{2-p} \bar{\theta} \varrho^p],$$

provided that $K_{4\varrho}(x_o) \times (t_o, t_o + b\xi^{2-p}\bar{\theta}(4\varrho)^p] \subset E_T$.

Remark 3.1. Once we have Proposition 3.1 at our disposal, the constant ξ can be chosen smaller if necessary. As such, a pointwise estimate for v can be actually claimed over longer cylinders as long as they are inside E_T .

3.1. Preliminary estimates. Before coming to the actual proof of Proposition 3.1, we present some structural estimates for non-negative super-solutions to (3.1) under the structure assumptions (1.3). The following result has appeared in [14], to which we refer for the proof.

Lemma 3.2. *Consider a non-negative, local weak super-solution v to the parabolic equation (3.1), with structure conditions (1.3) in E_T . Assume that $K_{2\varrho}(y) \times [\bar{t}, \bar{t} + S] \subset E_T$ and that*

$$\inf_{K_{2\varrho}(y)} v(\cdot, \bar{t}) \geq k, \quad \text{for some } k > 0.$$

Then, there exists $\kappa \in (0, 1)$, depending only on the data $\{N, p, C_o, C_1\}$, such that for all $t \in (\bar{t}, \bar{t} + S]$ we have

$$\inf_{K_{\varrho}(y)} v(\cdot, t) \geq \frac{k}{2} \left(1 + \frac{t - \bar{t}}{\kappa k^{2-p}(2\varrho)^p} \right)^{\frac{1}{2-p}}.$$

We also need the following weak Harnack inequality, proved in [16]; see also [10, Chapter 5, Section 7].

Theorem 3.1. *Consider a non-negative, local weak super-solution v to the parabolic equation (3.1), with structure conditions (1.3) in E_T . Assume that $K_{16\varrho}(x_o) \subset E$. There exist constants $c, \gamma > 1$, depending only on the data $\{N, p, C_o, C_1\}$, such that, for a.e. $s \in (0, T)$,*

$$\int_{K_{\varrho}(x_o)} v(x, s) \, dx \leq c \left(\frac{\varrho^p}{T - s} \right)^{\frac{1}{p-2}} + \gamma \inf_{K_{4\varrho}(x_o)} v(\cdot, t), \quad (3.2)$$

for all times

$$s + \frac{1}{2}\theta\varrho^p \leq t \leq s + 2\theta\varrho^p,$$

where

$$\theta = \min \left\{ c^{2-p} \frac{T - s}{\varrho^p}, \left[\int_{K_{\varrho}(x_o)} v(x, s) \, dx \right]^{2-p} \right\}. \quad (3.3)$$

Remark 3.2. In practice, it will be useful to choose s and ϱ to satisfy

$$s + 2c^{p-2} \left[\int_{K_{\varrho}(x_o)} v(x, s) \, dx \right]^{2-p} \varrho^p < T, \quad (3.4)$$

that is,

$$c \left(\frac{\varrho^p}{T - s} \right)^{\frac{1}{p-2}} < \frac{1}{2^{\frac{1}{p-2}}} \int_{K_{\varrho}(x_o)} v(x, s) \, dx;$$

then, according to (3.3), we have

$$\theta = \left[\int_{K_\varrho(x_o)} v(x, s) \, dx \right]^{2-p},$$

and also by (3.2),

$$\int_{K_\varrho(x_o)} v(x, s) \, dx \leq \gamma_H \inf_{K_{4\varrho}(x_o)} v(\cdot, t), \quad (3.5)$$

for all times

$$s + \frac{1}{2}\theta\varrho^p \leq t \leq s + 2\theta\varrho^p.$$

Moreover, $\gamma_H = \gamma/(1 - 2^{2-p})$, and therefore, the constant is stable as $p \downarrow 2$.

We can show the following estimate by relying on the previous results and working as in the proof of Lemma 3.2 of [12].

Lemma 3.3. *Consider a non-negative, local weak super-solution v to the parabolic equation (3.1), with structure conditions (1.3) in E_T . For $k > 0$, define*

$$v_k \stackrel{\text{def}}{=} k - (k - v)_+,$$

introduce three nested cylinders

$$\begin{cases} \mathbf{Q}_2 = K_\varrho(x_o) \times (t_o + \frac{1}{4}\bar{\theta}\varrho^p, t_o + \frac{3}{4}\bar{\theta}\varrho^p], \\ \mathbf{Q}_1 = K_{\frac{3}{2}\varrho}(x_o) \times (t_o + \frac{1}{8}\bar{\theta}\varrho^p, t_o + \frac{7}{8}\bar{\theta}\varrho^p], \\ \mathbf{Q}_o = K_{2\varrho}(x_o) \times (t_o, t_o + \bar{\theta}\varrho^p], \end{cases}$$

with a parameter $\bar{\theta} > 0$, and assume that $\mathbf{Q}_o \subset E_T$. For every $\eta \in (0, 1)$, there exists a constant $\tilde{\gamma} > 1$, depending only on the data $\{N, p, C_o, C_1\}$ and η , such that

$$\int_{t_o + \frac{1}{4}\bar{\theta}\varrho^p}^{t_o + \frac{3}{4}\bar{\theta}\varrho^p} \int_{K_{\frac{3}{2}\varrho}(x_o)} |D(v_k \zeta)|^p \, dx dt \leq k\varrho^{N-p} \mathbb{I}, \quad (3.6)$$

where

$$\mathbb{I} \stackrel{\text{def}}{=} \tilde{\gamma} \left[\sup_{t_o < t < t_o + \bar{\theta}\varrho^p} \int_{K_{2\varrho}(x_o)} v_k(x, t) \, dx \right]^{p-1} + \frac{\eta}{\bar{\theta}^{\frac{p-1}{p-2}}},$$

and $\zeta \in C_o^1(\mathbf{Q}_1; [0, 1])$ satisfies $\zeta = 1$ in \mathbf{Q}_2 , $|D\zeta| \leq 4/\varrho$ and $|\partial_t \zeta| \leq 16/(\bar{\theta}\varrho^p)$.

Proof. According to [10, Chapter 3, Lemma 1.1], v_k is a local weak super-solution to (3.1), with (1.3) in E_T . We use the testing function $(k - v_k)\zeta^p$ in the weak formulation of v_k , modulo a standard Steklov average. After standard calculations, we get the following energy estimate:

$$\begin{aligned} \iint_{\mathbf{Q}_1} |D(v_k \zeta)|^p \, dx dt &\leq \gamma k \iint_{\mathbf{Q}_1} |Dv_k|^{p-1} |D\zeta| \, dx dt + \gamma \iint_{\mathbf{Q}_1} v_k^p |D\zeta|^p \, dx dt \\ &\quad + \gamma k \iint_{\mathbf{Q}_1} v_k |\partial_t \zeta| \, dx dt, \end{aligned}$$

for some $\gamma = \gamma(p, C_o, C_1)$. This is precisely what one obtains in the proof of [12, Lemma 3.2]. Starting from this, one can reproduce the arguments and arrive at

$$\iint_{\mathbf{Q}_1} |D(v_k \zeta)|^p dx dt \leq C_\eta k \bar{\theta} \varrho^N \left[\sup_{t_o < t < t_o + \bar{\theta} \varrho^p} \int_{K_{2\varrho}(x_o)} v_k(x, t) dx \right]^{p-1} + k \bar{\theta} \varrho^N \frac{\eta}{\bar{\theta}^{\frac{p-1}{p-2}}},$$

for any $\eta \in (0, \frac{1}{2})$ and some $C_\eta > 1$. A consequence of the previous estimate is

$$\begin{aligned} & \int_{t_o + \frac{1}{4}\bar{\theta}\varrho^p}^{t_o + \frac{3}{4}\bar{\theta}\varrho^p} \int_{K_{\frac{3}{2}\varrho}(x_o)} |D(v_k \zeta)|^p dx dt \\ & \leq \frac{1}{\frac{1}{2}\bar{\theta}\varrho^p} \int_{t_o + \frac{1}{8}\bar{\theta}\varrho^p}^{t_o + \frac{7}{8}\bar{\theta}\varrho^p} \int_{K_{\frac{3}{2}\varrho}(x_o)} |D(v_k \zeta)|^p dx dt \\ & \leq 2C_\eta k \varrho^{N-p} \left[\sup_{t_o < t < t_o + \bar{\theta}\varrho^p} \int_{K_{2\varrho}(x_o)} v_k(x, t) dx \right]^{p-1} + k \varrho^{N-p} \frac{2\eta}{\bar{\theta}^{\frac{p-1}{p-2}}}. \end{aligned}$$

The proof is concluded by redefining 2η as η and choosing $\tilde{\gamma} = 2C_\eta$. \square

Remark 3.3. The role of $\bar{\theta}$ becomes significant when it restores the homogeneity of the estimate (3.6). In fact, if we let

$$\bar{\theta} = (\delta k)^{2-p},$$

for some $\delta \in (0, 1)$ to be determined later, it is straightforward to see that

$$\frac{1}{\bar{\theta}^{\frac{p-1}{p-2}}} = (\delta k)^{p-1} < k^{p-1}.$$

Moreover, by definition of v_k , we have

$$\sup_{t_o < t < t_o + \bar{\theta}\varrho^p} \int_{K_{2\varrho}(x_o)} v_k(x, t) dx \leq k.$$

Therefore, under the choice of $\bar{\theta}$, we have $\mathbb{I} \leq 2\tilde{\gamma}k^{p-1}$. Keep in mind that $\tilde{\gamma} = \tilde{\gamma}(\eta)$.

3.2. Proof of Proposition 3.1. For some $\varepsilon \in (0, \frac{1}{2})$ to be determined later, without loss of generality, we can assume that $\frac{1}{2}\varepsilon^{-p}$ is an integer; hence, we divide the interval $(t_o + \frac{1}{4}\bar{\theta}\varrho^p, t_o + \frac{3}{4}\bar{\theta}\varrho^p]$ into $\frac{1}{2}\varepsilon^{-p}$ sub-intervals, each of length $\bar{\theta}(\varepsilon\varrho)^p$. In the present subsection, we let $\bar{\theta} = (\delta k)^{2-p}$, for some $\delta \in (0, 1)$, so that Remark 3.3 applies. The actual form of δ will become clear in the course of the proof.

By (3.6), there must exist one of these sub-intervals, say $(t_*, t_* + \bar{\theta}(\varepsilon\varrho)^p]$, such that

$$\int_{t_*}^{t_* + \bar{\theta}(\varepsilon\varrho)^p} \int_{K_{\frac{3}{2}\varrho}(x_o)} |D(v_k \zeta)|^p dx dt \leq k \varrho^{N-p} \mathbb{I}, \quad (3.7)$$

which immediately implies the following estimate:

$$\int_{t_*}^{t_* + \frac{1}{2}\bar{\theta}(\varepsilon\varrho)^p} \int_{K_{\frac{3}{2}\varrho}(x_o)} |D(v_k \zeta)|^p dx dt \leq 2k \varrho^{N-p} \mathbb{I}. \quad (3.8)$$

Note that these estimates do not require any quantitative knowledge of ε , which is still to be selected. From (3.8), recalling that $\zeta \equiv 1$ in $\mathbf{Q}_2 = K_\varrho(x_o) \times (t_o + \frac{1}{4}\bar{\theta}\varrho^p, t_o + \frac{3}{4}\bar{\theta}\varrho^p]$, and also that $\mathbb{I} \leq 2\tilde{\gamma}k^{p-1}$, $\tilde{\gamma} = \tilde{\gamma}(\eta)$ from Remark 3.3, we infer that there exists some

$$\bar{t} \in [t_*, t_* + \frac{1}{2}\bar{\theta}(\varepsilon\varrho)^p]$$

satisfying

$$\int_{K_\varrho(x_o) \times \{\bar{t}\}} |Dv_k|^p dx dt \leq 4\tilde{\gamma}k^p \varrho^{N-p},$$

while the measure-theoretical information yields

$$|[v_k(\cdot, \bar{t}) = k] \cap K_\varrho(x_o)| \geq \alpha |K_\varrho|.$$

The last two estimates permit us to apply Lemma 2.1. Indeed, using the lemma for $v_k(\cdot, \bar{t})$, with $\delta = \lambda = \frac{1}{2}$, we find

$$\varepsilon = \bar{c}\alpha^2 \tag{3.9}$$

as in (2.3), with $\bar{c} \in (0, \frac{1}{2})$ depending only on the data $\{N, p, C_o, C_1\}$ and η , such that

$$|[v_k(\cdot, \bar{t}) > \frac{1}{2}k] \cap K_{\varepsilon\varrho}(y)| > \frac{1}{2}|K_{\varepsilon\varrho}|, \tag{3.10}$$

for some $K_{\varepsilon\varrho}(y) \subset K_\varrho(x_o)$.

Since v_k is a weak super-solution to (3.1), with structure conditions (1.3) in E_T , we apply the weak Harnack inequality (3.5) to v_k on $K_{\varepsilon\varrho}(y)$ at time \bar{t} . As a result, we obtain

$$\int_{K_{\varepsilon\varrho}(y)} v_k(x, \bar{t}) dx \leq \gamma_H \inf_{K_{4\varepsilon\varrho}(y)} v_k(\cdot, t), \tag{3.11}$$

for all times

$$\bar{t} + \frac{1}{2}\theta(\varepsilon\varrho)^p \leq t \leq \bar{t} + 2\theta(\varepsilon\varrho)^p, \quad \text{where} \quad \theta = \left[\int_{K_{\varepsilon\varrho}(y)} v_k(x, \bar{t}) dx \right]^{2-p},$$

provided that (3.4) holds, that is, $\bar{t} + 2c^{p-2}\theta(\varepsilon\varrho)^p < T$ in the current set-up.

Let us take a closer look at the above conclusion using (3.10). Notice first that

$$\int_{K_{\varepsilon\varrho}(y)} v_k(x, \bar{t}) dx \geq \frac{1}{|K_{\varepsilon\varrho}|} \int_{K_{\varepsilon\varrho}(y) \cap [v_k > \frac{1}{2}k]} v_k(x, \bar{t}) dx > \frac{1}{4}k,$$

and, on the other hand, by definition of v_k , we have

$$\int_{K_{\varepsilon\varrho}(y)} v_k(x, \bar{t}) dx \leq k.$$

Hence, we can estimate θ as

$$k^{2-p} \leq \theta = \left[\int_{K_{\varepsilon\varrho}(y)} v_k(x, \bar{t}) dx \right]^{2-p} < (\frac{1}{4}k)^{2-p}, \tag{3.12}$$

and, from (3.11), we infer that

$$v_k \geq \frac{k}{4\gamma_H} \stackrel{\text{def}}{=} \bar{\eta}k, \quad \text{a.e. in } K_{\varepsilon\varrho}(y) \times [\bar{t} + \frac{1}{2}\theta(\varepsilon\varrho)^p, \bar{t} + 2\theta(\varepsilon\varrho)^p]. \tag{3.13}$$

By (3.12), $\bar{t} \leq t_o + \bar{\theta}\varrho^p$, and $\bar{\theta} = (\delta k)^{2-p}$, the requirement $\bar{t} + 2c^{p-2}\theta(\varepsilon\varrho)^p < T$ is fulfilled if

$$t_o + 2^p c^{p-2} \bar{\theta} \varrho^p < T. \quad (3.14)$$

No matter the value of $\bar{t} \in [t_*, t_* + \frac{1}{2}\bar{\theta}(\varepsilon\varrho)^p]$, provided δ in $\theta = (\delta k)^{2-p}$ is taken sufficiently small, we can surely conclude that

$$K_{\varepsilon\varrho}(y) \times [\bar{t} + \frac{1}{2}\theta(\varepsilon\varrho)^p, \bar{t} + 2\theta(\varepsilon\varrho)^p] \subset K_{\varrho}(x_o) \times (t_*, t_* + \bar{\theta}(\varepsilon\varrho)^p]. \quad (3.15)$$

Indeed, since

$$\bar{t} + 2\theta(\varepsilon\varrho)^p \leq t_* + \frac{1}{2}\bar{\theta}(\varepsilon\varrho)^p + 2\theta(\varepsilon\varrho)^p,$$

in order to have the wanted inclusion, it suffices to require

$$t_* + \frac{1}{2}\bar{\theta}(\varepsilon\varrho)^p + 2\theta(\varepsilon\varrho)^p \leq t_* + \bar{\theta}(\varepsilon\varrho)^p, \quad i.e., \quad \theta \leq \frac{1}{4}\bar{\theta},$$

which is implied if

$$\delta \leq 4^{\frac{p-1}{2-p}}, \quad (3.16)$$

using $\bar{\theta} = (\delta k)^{2-p}$ and the right-hand side of (3.12). This sets a smallness requirement on δ , while a more precise choice of δ is yet to come.

The inclusion (3.15) allows us to estimate

$$\int_{t_*}^{t_* + \bar{\theta}(\varepsilon\varrho)^p} \int_{K_{\frac{3}{2}\varrho}(x_o)} |D(v_k \zeta)|^p dx dt \geq \frac{\frac{3}{2}\theta(\varepsilon\varrho)^p}{\bar{\theta}(\varepsilon\varrho)^p} \int_{\bar{t} + \frac{1}{2}\theta(\varepsilon\varrho)^p}^{\bar{t} + 2\theta(\varepsilon\varrho)^p} \int_{K_{\frac{3}{2}\varrho}(x_o)} |D(v_k \zeta)|^p dx dt,$$

and hence, by (3.7), we have

$$\frac{3\theta}{2\bar{\theta}} \int_{\bar{t} + \frac{1}{2}\theta(\varepsilon\varrho)^p}^{\bar{t} + 2\theta(\varepsilon\varrho)^p} \int_{K_{\frac{3}{2}\varrho}(x_o)} |D(v_k \zeta)|^p dx dt \leq k\varrho^{N-p} \mathbb{I}.$$

The last display, together with (3.13), gives some $\tilde{t} \in (\bar{t} + \frac{1}{2}\theta(\varepsilon\varrho)^p, \bar{t} + 2\theta(\varepsilon\varrho)^p]$, such that

$$\frac{3\theta}{2\bar{\theta}} \int_{K_{\frac{3}{2}\varrho}(x_o) \times \{\tilde{t}\}} |D(v_k \zeta)|^p dx \leq k\varrho^{N-p} \mathbb{I} \quad \text{and} \quad \inf_{K_{\varepsilon\varrho}(y)} v_k(\cdot, \tilde{t}) \geq \bar{\eta}k.$$

By properties of the p -capacity (see Appendix A), $K_{\varepsilon\varrho}(y) \subset K_{\varrho}(x_o)$, and $K_{\frac{3}{2}\varrho}(x_o) \subset K_{3\varrho}(y)$, we obtain

$$k\varrho^{N-p} \mathbb{I} \geq \frac{3\theta}{2\bar{\theta}} (\bar{\eta}k)^p \text{cap}_p(K_{\varepsilon\varrho}(y), K_{\frac{3}{2}\varrho}(x_o)) \geq \frac{3\theta}{2\bar{\theta}} (\bar{\eta}k)^p \text{cap}_p(K_{\varepsilon\varrho}(y), K_{3\varrho}(y)).$$

By $\bar{\theta} = (\delta k)^{2-p}$ and the estimate of θ in (3.12), we further compute

$$\begin{aligned} k\varrho^{N-p} \mathbb{I} &\geq \frac{3\delta^{p-2}}{2} (\bar{\eta}k)^p \text{cap}_p(K_{\varepsilon\varrho}(y), K_{3\varrho}(y)), \quad i.e., \\ \mathbb{I} &\geq \frac{3\bar{\eta}^p}{2} \delta^{p-2} \frac{\text{cap}_p(K_{\varepsilon\varrho}(y), K_{3\varrho}(y))}{\varrho^{N-p}} k^{p-1}. \end{aligned}$$

Recalling the definition of \mathbb{I} and also the scaling property (A.1) of p -capacity, the last inequality yields

$$\frac{3\bar{\eta}^p}{2} \delta^{p-2} \text{cap}_p(K_{\varepsilon}, K_3) k^{p-1} \leq \tilde{\gamma} \left[\sup_{t_o < t < t_o + \bar{\theta}\varrho^p} \int_{K_{2\varrho}(x_o)} v_k(x, t) dx \right]^{p-1} + \eta\delta^{p-1} k^{p-1}.$$

At this point, we select the parameters δ and η so that the second term on the right-hand side will be absorbed into the left. This leads to the choices

$$\delta(\varepsilon) = \bar{\delta} \operatorname{cap}_p(K_\varepsilon, K_3), \quad \eta = \frac{3\bar{\eta}^p}{4\bar{\delta}}. \quad (3.17)$$

Here, $\bar{\delta} \in (0, 1)$ can be selected in terms of $\{N, p\}$ only to guarantee the previous smallness requirement (3.16) on δ . In fact, by (A.3), we have $\operatorname{cap}_p(K_\varepsilon, K_3) \leq \gamma(N, p)$, and hence it suffices to take

$$\bar{\delta} = \frac{4^{\frac{p-1}{2-p}}}{\gamma},$$

whereas $\bar{\eta} = 1/(4\gamma_H)$ is from (3.13). Thus, at this point, η has been determined in (3.17) by the data $\{N, p, C_o, C_1\}$, and also ε in (3.9). Consequently, these choices lead to

$$\frac{3\bar{\eta}^p \bar{\delta}^{p-2}}{4\bar{\gamma}} [\operatorname{cap}_p(K_\varepsilon, K_3)]^{p-1} k^{p-1} \leq \left[\sup_{t_o < t < t_o + \bar{\theta} \varrho^p} \int_{K_{2\varrho}(x_o)} v_k(x, t) \, dx \right]^{p-1},$$

i.e.,

$$\frac{\eta \bar{\delta}^{p-1}}{\bar{\gamma}} [\operatorname{cap}_p(K_\varepsilon, K_3)]^{p-1} k^{p-1} \leq \left[\sup_{t_o < t < t_o + \bar{\theta} \varrho^p} \int_{K_{2\varrho}(x_o)} v_k(x, t) \, dx \right]^{p-1};$$

in other words, we have

$$\eta_* \bar{\delta} \operatorname{cap}_p(K_\varepsilon, K_3) k \leq \sup_{t_o < t < t_o + \bar{\theta} \varrho^p} \int_{K_{2\varrho}(x_o)} v_k(x, t) \, dx, \quad (3.18)$$

where

$$\bar{\theta} = (\delta k)^{2-p} = [\operatorname{cap}_p(K_\varepsilon, K_3)]^{2-p} (\bar{\delta} k)^{2-p} \quad \text{and} \quad \eta_* = \left(\frac{\eta}{\bar{\gamma}} \right)^{\frac{1}{p-1}}. \quad (3.19)$$

Note that $\bar{\eta} = 1/(4\gamma_H)$ from (3.13), η is given by (3.17), and $\tilde{\gamma} = \tilde{\gamma}(\eta)$ is fixed once η is chosen in (3.17). Hence, $\eta_* \in (0, 1)$ depends only on the data $\{N, p, C_o, C_1\}$.

Now, we analyze the consequences of (3.18), together with the weak Harnack inequality. Recall the choice of θ in (3.19), and for simplicity, denote

$$\tilde{\delta}(\varepsilon) \stackrel{\text{def}}{=} \operatorname{cap}_p(K_\varepsilon, K_3) \quad \Rightarrow \quad \delta(\varepsilon) = \tilde{\delta} \tilde{\delta}(\varepsilon), \quad \bar{\theta} = [\tilde{\delta} \tilde{\delta}(\varepsilon) k]^{2-p}.$$

Let $t_1 \in [t_o, t_o + \bar{\theta} \varrho^p]$ be the level where the supremum in (3.18) is attained, so that

$$\eta_* \delta(\varepsilon) k \leq \int_{K_{2\varrho}(x_o)} v_k(x, t_1) \, dx. \quad (3.20)$$

The weak Harnack inequality (3.5) applied to v_k on $K_{2\varrho}(x_o)$ at time t_1 yields

$$\int_{K_{2\varrho}(x_o)} v_k(x, t_1) \, dx \leq \gamma_H \inf_{K_{4\varrho}(x_o)} v_k(\cdot, t), \quad (3.21)$$

for all times

$$t_1 + \frac{1}{2}\theta_1(2\varrho)^p \leq t \leq t_1 + 2\theta_1(2\varrho)^p, \quad \text{where} \quad \theta_1 = \left[\int_{K_{2\varrho}(x_o)} v_k(x, t_1) \, dx \right]^{2-p},$$

provided that (3.4) holds, *i.e.*, in the current set-up,

$$t_1 + 2c^{p-2}\theta_1(2\varrho)^p < T. \quad (3.22)$$

Assume that (3.22) holds for the moment; then, combining (3.20) and (3.21), we obtain

$$\inf_{K_{4\varrho}(x_o)} v_k(\cdot, \hat{t}) \geq \frac{\eta_*}{\gamma_H} \delta(\varepsilon) k, \quad \text{where } \hat{t} \stackrel{\text{def}}{=} t_1 + 2\theta_1(2\varrho)^p. \quad (3.23)$$

Now, let us examine what is required for (3.22) to hold true. From (3.20), it is apparent that

$$\theta_1 = \left[\int_{K_{2\varrho}(x_o)} v_k(x, t_1) dx \right]^{2-p} \leq (\eta_* \delta(\varepsilon) k)^{2-p} = \eta_*^{2-p} \bar{\theta}.$$

Recalling that $t_1 \in [t_o, t_o + \bar{\theta}\varrho^p]$, we estimate

$$t_1 + 2c^{p-2}\theta_1(2\varrho)^p \leq t_o + \bar{\theta}\varrho^p + 2c^{p-2}\eta_*^{2-p}\bar{\theta}(2\varrho)^p = t_o + (1 + 2^{p+1}c^{p-2}\eta_*^{2-p})\bar{\theta}\varrho^p.$$

Hence, condition (3.22) amounts to requiring

$$t_o + (1 + 2^{p+1}c^{p-2}\eta_*^{2-p})\bar{\theta}\varrho^p < T, \quad (3.24)$$

which we may assume. Likewise, we also have

$$\hat{t} = t_1 + 2\theta_1(2\varrho)^p \leq t_o + \bar{\theta}\varrho^p + 2^{p+1}\eta_*^{2-p}\bar{\theta}\varrho^p = t_o + (1 + 2^{p+1}\eta_*^{2-p})\bar{\theta}\varrho^p. \quad (3.25)$$

Finally, we aim to use (3.23) and apply Lemma 3.2. Indeed, consider $t \in (t_o + 4^p\eta_*^{2-p}\bar{\theta}\varrho^p, t_o + 8^p\eta_*^{2-p}\bar{\theta}\varrho^p]$. Then, estimate (3.25) gives that $0 < t - \hat{t} \leq 8^p\eta_*^{2-p}\bar{\theta}\varrho^p$, and also that

$$1 + \frac{t - \hat{t}}{\kappa \left[\frac{\eta_*}{\gamma_H} \delta(\varepsilon) k \right]^{2-p} (4\varrho)^p} \leq 1 + \frac{2^p}{\kappa \gamma_H^{p-2}}.$$

Hence, a straightforward application of Lemma 3.2 yields

$$\inf_{K_\varrho(x_o)} v_k(\cdot, t) \geq \frac{\eta_* \delta(\varepsilon) k}{2\gamma_H} \left(1 + \frac{2^p}{\kappa \gamma_H^{p-2}} \right)^{\frac{1}{2-p}} \stackrel{\text{def}}{=} \tilde{\eta}[\eta_* \delta(\varepsilon) k], \quad (3.26)$$

for all $t \in (t_o + 4^p[\eta_* \delta(\varepsilon) k]^{2-p}\varrho^p, t_o + 8^p[\eta_* \delta(\varepsilon) k]^{2-p}\varrho^p]$, provided that

$$t_o + 8^p\eta_*^{2-p}\bar{\theta}\varrho^p < T. \quad (3.27)$$

Now, it suffices to define $\xi = \tilde{\eta}\eta_*$ and choose $b = \max\{4^p\tilde{\eta}^{p-2}, 4^p(\tilde{\eta}c)^{p-2}\}$. In this way the desired pointwise estimate is given in (3.26) if we require $t_o + b\xi^{2-p}\bar{\theta}(4\varrho)^p < T$, so that the conditions (3.14), (3.22), (3.24) and (3.27) are satisfied. Taking into account the definitions of ε in (3.9) and $\delta(\varepsilon)$ in (3.17), we conclude.

4. PROOF OF THEOREM 1.1

Let u_ϵ be a solution to the regularized problem (2.2); we recall that u_ϵ is locally Hölder continuous, hence, in particular, defined everywhere. From now on, we will omit the ϵ for simplicity. Moreover, we will focus primarily on the case $\underline{p} = N$, discussing the changes needed to cover the remaining cases at the end of the section.

Let $b > 1$, \bar{c} , $\bar{\delta}$, $\xi \in (0, 1)$ be the constants stipulated in Proposition 3.1; they depend only on $\{N, C_o, C_1\}$, whereas the constant $c_1 \in (0, 1)$ postulated in Lemma 2.4 depends on $\{\nu, N, C_o, C_1\}$. Suppose there are parameters $\varrho > 0$ and $\omega \in (0, 1)$ such that

$$Q_{8\varrho}(b\xi^{2-N}\bar{\theta}) \subset E_T, \quad \underset{Q_{8\varrho}(b\xi^{2-N}\bar{\theta})}{\text{osc}} u \leq \omega \quad (4.1)$$

where

$$\bar{\theta} \stackrel{\text{def}}{=} \left[\left(\frac{1}{8}\omega \right) \bar{\delta} \text{cap}_N(K_{\bar{c}\alpha^2}, K_3) \right]^{2-N}, \quad \alpha \stackrel{\text{def}}{=} c_1 \left(\frac{1}{8}\omega \right)^3. \quad (4.2)$$

Here and in what follows, we assume up to a translation that $(x_o, t_o) = (0, 0)$. Moreover, we consider the quantities

$$\mu^+ = \sup_{Q_{8\varrho}(b\xi^{2-N}\bar{\theta})} u, \quad \mu^- = \inf_{Q_{8\varrho}(b\xi^{2-N}\bar{\theta})} u,$$

and assume that

$$\mu^+ - \mu^- \geq \frac{1}{2}\omega,$$

dealing with the opposite situation at the end. Due to this assumption, we have that one of the following **two cases** must hold: either

$$\mu^+ - \frac{1}{8}\omega \geq \frac{1}{8}\omega \quad (4.3)$$

or

$$\mu^- + \frac{1}{8}\omega \leq -\frac{1}{8}\omega. \quad (4.4)$$

Let us first suppose that the **first case** (4.3) holds true. We perform the reduction of oscillation in the following two subsections under the first case. For that, we let

$$I \stackrel{\text{def}}{=} (-2b\xi^{2-N}\bar{\theta}\varrho^N, -b\xi^{2-N}\bar{\theta}\varrho^N],$$

and further consider the following **two alternatives**:

$$\exists \bar{t} \in I \quad \text{such that} \quad |[u(\cdot, \bar{t}) \leq \mu^- + \frac{1}{8}\omega] \cap K_\varrho| \leq c_1 \left(\frac{1}{8}\omega \right)^3 |K_\varrho|, \quad (4.5)$$

$$\forall t \in I \quad \text{we have} \quad |[u(\cdot, t) \leq \mu^- + \frac{1}{8}\omega] \cap K_\varrho| > \underbrace{c_1 \left(\frac{1}{8}\omega \right)^3}_{=\alpha} |K_\varrho|. \quad (4.6)$$

4.1. First case – first alternative. In this section, we assume that the **first case** (4.3) holds true and consider the **first alternative** (4.5). Then, Lemma 2.4 for super-solutions yields

$$u\left(\cdot, \bar{t} + \frac{\delta}{2^N} \left(\frac{1}{8}\omega \right)^{2-N} \varrho^N\right) \geq \mu^- + \frac{1}{32}\omega \quad \text{in } K_{\frac{1}{4}\varrho},$$

where $\delta \in (0, 1)$ is the quantity stipulated in Lemma 2.4, which depends only on the data $\{\nu, N, C_o, C_1\}$. Then, for any $\tilde{\xi} \in (0, \frac{1}{32}]$ we obviously have

$$u\left(\cdot, \bar{t} + \frac{\delta}{2^N} \left(\frac{1}{8}\omega \right)^{2-N} \varrho^N\right) \geq \mu^- + \tilde{\xi}\omega \quad \text{in } K_{\frac{1}{4}\varrho}.$$

We want to apply Lemma 2.3 for super-solutions, relying on such pointwise information with time level $t_1 = \bar{t} + \frac{\delta}{2^N} \left(\frac{1}{8}\omega \right)^{2-N} \varrho^N$ for a proper $\tilde{\xi}$ to be chosen; provided we let $\theta = \gamma_o(\tilde{\xi}\omega)^{2-N}$ in Lemma 2.3, the measure-theoretical condition of Lemma 2.3 is automatically satisfied, and we conclude that

$$u \geq \mu^- + \frac{1}{2}\tilde{\xi}\omega \quad (4.7)$$

in

$$K_{\frac{1}{8}\varrho} \times \left(\bar{t} + \frac{\delta}{2^N} \left(\frac{1}{8}\omega\right)^{2-N} \varrho^N, \bar{t} + \frac{\delta}{2^N} \left(\frac{1}{8}\omega\right)^{2-N} \varrho^N + \gamma_o(\tilde{\xi}\omega)^{2-N} \frac{1}{4^N} \varrho^N \right], \quad (4.8)$$

where $\tilde{\xi}$ is still to be determined.

Considering the location of \bar{t} in I , and taking into account the definition of $\bar{\theta}$ in (4.2), in order for the previous estimate (4.7) to reach $t = 0$, it suffices to choose $\tilde{\xi}$ such that

$$\begin{aligned} \frac{\delta}{2^N} \left(\frac{1}{8}\omega\right)^{2-N} \varrho^N + \gamma_o(\tilde{\xi}\omega)^{2-N} \frac{1}{4^N} \varrho^N &\geq 2b\xi^{2-N} \bar{\theta} \varrho^N \\ &= 2b\xi^{2-N} \left[\left(\frac{1}{8}\omega\right) \bar{\delta} \text{cap}_N(K_{\bar{c}\alpha^2}, K_3) \right]^{2-N} \varrho^N, \end{aligned} \quad (4.9)$$

which is the case if, discarding the term containing δ ,

$$\tilde{\xi} \leq \left(\frac{\gamma_o}{b} \right)^{\frac{1}{N-2}} 2^{\frac{5(1-N)}{N-2}} \xi \bar{\delta} [\text{cap}_N(K_{\bar{c}\alpha^2}, K_3)].$$

By the capacity estimate of (A.4)₂, and the definition of α in (4.2), we calculate

$$\text{cap}_N(K_{\bar{c}\alpha^2}, K_3) = b_2 |\ln[\bar{c}c_1^2 \left(\frac{1}{8}\omega\right)^6]|^{-(N-1)} = 6^{-(N-1)} b_2 |\ln(\bar{\xi}\omega)|^{-(N-1)}, \quad (4.10)$$

for some b_2 depending only on N , where we denoted

$$\bar{\xi}^6 \stackrel{\text{def}}{=} \frac{\bar{c}c_1^2}{8^6}$$

for simplicity. Note that $\bar{\xi} = \gamma(N, C_o, C_1)/\nu_*^2$ because of the dependencies of c_1 .

It is apparent that $\bar{\xi} \in (0, 1)$. Then, we can choose

$$\tilde{\xi} = \underbrace{\left(\frac{\gamma_o}{b} \right)^{\frac{1}{N-2}} 2^{\frac{5(1-N)}{N-2}} \xi \bar{\delta} 6^{-(N-1)} b_2 |\ln(\bar{\xi}\omega)|^{-(N-1)}}_{\stackrel{\text{def}}{=} 2\bar{\eta}}. \quad (4.11)$$

Note that $\bar{\eta} = \bar{\eta}(N, C_o, C_1)$ is independent of ν . Next, by imposing $\xi < 2^{-18}$ (see Remark 3.1), we further estimate that

$$\begin{aligned} b\xi^{2-N} \bar{\theta} \varrho^N &= b\xi^{2-N} \left[\left(\frac{1}{8}\omega\right) \bar{\delta} \text{cap}_N(K_{\bar{c}\alpha^2}, K_3) \right]^{2-N} \varrho^N \\ &\stackrel{(4.10)}{=} b\xi^{2-N} \left[\left(\frac{1}{8}\omega\right) \bar{\delta} 6^{-(N-1)} b_2 |\ln(\bar{\xi}\omega)|^{-(N-1)} \right]^{2-N} \varrho^N \\ &> (32^N + 1) \theta \left(\frac{1}{8}\varrho\right)^N, \end{aligned}$$

where we defined

$$\theta \stackrel{\text{def}}{=} \left(\frac{\omega}{|\ln(\bar{\xi}\omega)|^{N-1}} \right)^{2-N}. \quad (4.12)$$

As a result of the last estimate, we have

$$-b\xi^{2-N} \bar{\theta} \varrho^N + \frac{\delta}{2^N} \left(\frac{1}{8}\omega\right)^{2-N} \varrho^N < -\theta \left(\frac{1}{8}\varrho\right)^N.$$

The above inequality, jointly with (4.9), guarantees that the cylinder (4.8) includes $Q_{\frac{1}{8}\varrho}(\theta)$, no matter where \bar{t} is in the interval I . Therefore, the pointwise estimate (4.7) can be claimed in $Q_{\frac{1}{8}\varrho}(\theta)$, and this, in turn, yields the oscillation estimate

$$\text{osc}_{Q_{\frac{1}{8}\varrho}(\theta)} u \leq (1 - \eta)\omega, \quad (4.13)$$

where θ is defined in (4.12) and

$$\eta \stackrel{\text{def}}{=} \frac{\bar{\eta}}{|\ln(\bar{\xi}\omega)|^{N-1}}. \quad (4.14)$$

Notice that $\bar{\eta}$ depends only on the data $\{N, C_o, C_1\}$ while $\bar{\xi}$ depends additionally on ν .

4.2. First case – second alternative. In this subsection, we still assume that the **first case** (4.3) holds true. However, we consider the **second alternative** (4.6). Since $\mu^+ - \frac{1}{8}\omega \geq \mu^- + \frac{1}{8}\omega$, we can rephrase it in the following way:

$$\forall t \in I \quad \text{we have} \quad |[u(\cdot, t) \leq \mu^+ - \frac{1}{8}\omega] \cap K_\varrho| > c_1(\frac{1}{8}\omega)^3 |K_\varrho|.$$

We introduce the new function

$$v \stackrel{\text{def}}{=} \frac{1}{8}\omega - (u - k)_+, \quad \text{with } k = \mu^+ - \frac{1}{8}\omega.$$

Since (4.3) ensures that $\mu^+ - \frac{1}{8}\omega > 0$, by Lemma 3.1, the above-defined v is a non-negative, weak super-solution to the parabolic equation (3.1), with structure conditions (1.3) in $Q_{8\varrho}(b\xi^{2-N}\bar{\theta})$. Moreover, in terms of v , the previous measure-theoretical information can be rephrased as

$$\forall t \in I, \quad |[v(\cdot, t) = \frac{1}{8}\omega] \cap K_\varrho| > \underbrace{c_1(\frac{1}{8}\omega)^3}_{=\alpha} |K_\varrho|.$$

In particular, such information holds for any $t \in J$, where

$$J \stackrel{\text{def}}{=} (-2b\xi^{2-N}\bar{\theta}\varrho^N, -(2b\xi^{2-N} - 1)\bar{\theta}\varrho^N],$$

and $\bar{\theta}$ is defined in (4.2). We can then apply Proposition 3.1, with $t_o = -2b\xi^{2-N}\bar{\theta}\varrho^N$ and $k = \frac{1}{8}\omega$, and conclude that

$$v \geq \frac{1}{8}\omega \xi \text{cap}_N(K_{\bar{c}\alpha^2}, K_3) \stackrel{(4.10)}{=} \frac{1}{8}\omega \xi 6^{-(N-1)} b_2 |\ln(\bar{\xi}\omega)|^{-(N-1)}$$

in $K_{\frac{1}{2}\varrho} \times (-b\xi^{2-N}\bar{\theta}\varrho^N, 0]$. Taking into account the definition of v , this yields the following reduction of oscillation:

$$\text{osc}_{Q_{\frac{1}{2}\varrho}(\theta)} u \leq (1 - \eta)\omega, \quad (4.15)$$

where θ is as in (4.12) and

$$\eta \stackrel{\text{def}}{=} \underbrace{\frac{1}{8}\xi 6^{-(N-1)} b_2}_{\stackrel{\text{def}}{=} \bar{\eta}} |\ln(\bar{\xi}\omega)|^{-(N-1)} = \frac{\bar{\eta}}{|\ln(\bar{\xi}\omega)|^{N-1}}. \quad (4.16)$$

Notice that the above-defined $\bar{\eta}$ depends only on the data $\{N, C_o, C_1\}$ and might differ from the one in (4.14). However, we tacitly take the smaller of the two. Therefore, we conclude that when the **first case** (4.3) holds, we have

$$\text{osc}_{Q_{\frac{1}{8}\varrho}(\theta)} u \leq \left(1 - \frac{\bar{\eta}}{|\ln(\bar{\xi}\omega)|^{N-1}}\right)\omega, \quad \text{where } \theta = \left(\frac{\omega}{|\ln(\bar{\xi}\omega)|^{N-1}}\right)^{2-N}. \quad (4.17)$$

Now, we assume that the **second case** (4.4) is satisfied. In the following two subsections, we perform the oscillation reduction for this case. Although the approach is

very similar, we consider it in detail since now the regularization parameter ϵ plays a role.

To this end, we consider the following **two alternatives**:

$$\exists \bar{t} \in I \quad \text{such that} \quad |[u(\cdot, \bar{t}) \geq \mu^+ - \frac{1}{8}\omega] \cap K_\rho| \leq c_1(\frac{1}{8}\omega)^3 |K_\rho|, \quad (4.18)$$

$$\forall t \in I \quad \text{we have} \quad |[u(\cdot, t) \geq \mu^+ - \frac{1}{8}\omega] \cap K_\rho| > c_1(\frac{1}{8}\omega)^3 |K_\rho|. \quad (4.19)$$

4.3. Second case – first alternative. Under the **second case** (4.4), we suppose the **first alternative** (4.18) holds true, and we work precisely as we did in § 4.1 when we assumed that (4.5) is satisfied. In particular, relying first on Lemma 2.4 for sub-solutions, and then on Lemma 2.3 for sub-solutions, we obtain for any $\tilde{\xi} \in (0, \frac{1}{32}]$ that

$$u \leq \mu^+ - \frac{1}{2}\tilde{\xi}\omega$$

in

$$K_{\frac{1}{8}\rho} \times \left(\bar{t} + \frac{\delta}{2^N}(\frac{1}{8}\omega)^{2-N} \rho^N, \bar{t} + \frac{\delta}{2^N}(\frac{1}{8}\omega)^{2-N} \rho^N + \gamma_o(\tilde{\xi}\omega)^{2-N} \frac{1}{4^N} \rho^N \right],$$

which corresponds to (4.7) and (4.8). Afterwards, repeating almost verbatim the same computations, we choose $\tilde{\xi}$ as in (4.11), define θ as in (4.12), and conclude the same oscillation estimate as (4.13):

$$\operatorname{osc}_{Q_{\frac{1}{8}\rho}(\theta)} u \leq (1 - \eta)\omega,$$

with θ as in (4.12) and η as in (4.14).

4.4. Second case – second alternative. Still under the **second case** (4.4), we consider instead the **second alternative** (4.19). Since $\mu^- + \frac{1}{8}\omega \leq \mu^+ - \frac{1}{8}\omega$, we can rephrase (4.19) as

$$\forall t \in I \quad \text{we have} \quad |[u(\cdot, t) \geq \mu^- + \frac{1}{8}\omega] \cap K_\rho| > c_1(\frac{1}{8}\omega)^3 |K_\rho|.$$

We introduce the new function

$$v \stackrel{\text{def}}{=} \frac{1}{8}\omega - (u - k)_-, \quad \text{with } k = \mu^- + \frac{1}{8}\omega.$$

Since $\mu^- + \frac{1}{8}\omega < -\frac{1}{8}\omega < 0$, in order to apply Lemma 3.1, and ensure that v is a non-negative, weak super-solution to the parabolic equation (3.1), with structure conditions (1.3) in Q_o , we need to require that $k < -\epsilon$; it suffices to assume that $-\frac{1}{8}\omega < -\epsilon$, that is, $\epsilon < \frac{1}{8}\omega$. Once such a condition on ϵ is satisfied, we can proceed almost verbatim as in § 4.2, and conclude the same oscillation estimate as (4.15), *i.e.*,

$$\operatorname{osc}_{Q_{\frac{1}{2}\rho}(\theta)} u \leq (1 - \eta)\omega,$$

with θ as in (4.12) and η as in (4.16).

Hence, under (4.1) assumed at the beginning, we have eventually found that either $\omega < 8\epsilon$ or (4.17) holds true. Notice that (4.17) takes into account also the case when $\mu^+ - \mu^- < \frac{1}{2}\omega$.

4.5. Derivation of the modulus of continuity. We have all the tools we need to derive a quantitative modulus of continuity. Let us summarize what we have achieved so far. Under the assumption that (4.1) holds true for $\varrho > 0$, and $\omega \in (0, 1)$, we obtained that either

$$Q_{\frac{1}{8}e}^{\text{osc}} u \leq \left(1 - \frac{\bar{\eta}}{|\ln(\bar{\xi}\omega)|^{N-1}}\right)\omega, \quad \text{with } \theta = \left(\frac{\omega}{|\ln(\bar{\xi}\omega)|^{N-1}}\right)^{2-N},$$

where $\bar{\eta}$ depends only on the data $\{N, C_o, C_1\}$ and $\bar{\xi}$ depends additionally on ν , or $\omega < 8\epsilon$.

Now, we need to iterate the argument. For that, we let

$$\omega_o = \omega, \quad \varrho_o = 8\varrho, \quad \theta_o = \theta, \quad \bar{\theta}_o = \bar{\theta},$$

set

$$\begin{aligned} \omega_1 &\stackrel{\text{def}}{=} \left(1 - \frac{\bar{\eta}}{|\ln(\bar{\xi}\omega_o)|^{N-1}}\right)\omega_o, \\ \bar{\theta}_1 &\stackrel{\text{def}}{=} \left[\left(\frac{1}{8}\omega_1\right)\bar{\delta} \text{cap}_N \left(K_{\bar{c}\bar{c}_1^2(\frac{1}{8}\omega_1)^6}, K_3\right)\right]^{2-N}, \end{aligned}$$

where $\bar{\delta}$, \bar{c} , b , ξ are the quantities stipulated in Proposition 3.1 in terms of $\{N, C_o, C_1\}$, and $c_1 = c_1(\nu, N, C_o, C_1)$ is the quantity in (2.5), and seek ϱ_1 to verify the following set inclusion

$$Q_{8\varrho_1}(b\xi^{2-N}\bar{\theta}_1) \subseteq Q_{\frac{1}{64}\varrho_o}(\theta_o).$$

Without loss of generality, we may assume

$$\frac{1}{|\ln(\bar{\xi}\omega_o)|^{N-1}} \leq \frac{1}{2},$$

which yields $\omega_1 \geq \frac{1}{2}\omega_o$. Hence, by the definition of $\bar{\theta}_1$, we can estimate

$$8^N b \xi^{2-N} \bar{\theta}_1 \varrho_1^N \leq 8^N b \xi^{2-N} \left[\left(\frac{1}{16}\omega_o\right)\bar{\delta} \text{cap}_N \left(K_{\bar{c}\bar{c}_1^2(\frac{1}{16}\omega_o)^6}, K_3\right)\right]^{2-N} \varrho_1^N.$$

Consequently, for the set inclusion, we select ϱ_1 to satisfy

$$8^N b \xi^{2-N} \left[\left(\frac{1}{16}\omega_o\right)\bar{\delta} \text{cap}_N \left(K_{\bar{c}\bar{c}_1^2(\frac{1}{16}\omega_o)^6}, K_3\right)\right]^{2-N} \varrho_1^N \leq \omega_o^{2-N} |\ln(\bar{\xi}\omega_o)|^{(N-1)(N-2)} \left(\frac{1}{64}\varrho_o\right)^N.$$

To proceed, we rewrite the capacity term using (A.4)₂ and (4.10):

$$\begin{aligned} &\text{cap}_N \left(K_{\bar{c}\bar{c}_1^2(\frac{1}{16}\omega_o)^6}, K_3\right) \\ &= b_2 \left| \ln \left[\frac{\bar{c}\bar{c}_1^2}{2^6} \left(\frac{\omega_o}{8}\right)^6 \right] \right|^{-(N-1)} = b_2 \left| \ln \left(\frac{\bar{\xi}^6 \omega_o^6}{2^6} \right) \right|^{-(N-1)} \\ &= b_2 \left| \ln(\bar{\xi}^6 \omega_o^6) - \ln 2^6 \right|^{-(N-1)} = b_2 [6 |\ln(\bar{\xi}\omega_o)| + 6 \ln 2]^{-(N-1)} \\ &= b_2 6^{-(N-1)} [|\ln(\bar{\xi}\omega_o)| + \ln 2]^{-(N-1)}. \end{aligned}$$

Hence, plugging it into the last estimate yields

$$\begin{aligned} &8^N b \xi^{2-N} \left[\left(\frac{1}{16}\omega_o\right)\bar{\delta} 6^{-(N-1)} b_2\right]^{2-N} [|\ln(\bar{\xi}\omega_o)| + \ln 2]^{(N-1)(N-2)} \varrho_1^N \\ &\leq \omega_o^{2-N} |\ln(\bar{\xi}\omega_o)|^{(N-1)(N-2)} \frac{1}{64^N} \varrho_o^N. \end{aligned}$$

Without loss of generality, we may assume that

$$|\ln(\bar{\xi}\omega_o)| > \ln 2 \quad \Rightarrow \quad |\ln(\bar{\xi}\omega_o)| + \ln 2 < 2|\ln(\bar{\xi}\omega_o)|.$$

Then, the above estimate is satisfied if

$$8^N b\xi^{2-N} \left[\frac{1}{16} \bar{\delta} 6^{-(N-1)} b_2 \right]^{2-N} [2|\ln(\bar{\xi}\omega_o)|]^{(N-1)(N-2)} \varrho_1^N \leq \frac{1}{64^N} |\ln(\bar{\xi}\omega_o)|^{(N-1)(N-2)} \varrho_o^N.$$

Hence, it suffices to choose ϱ_1 such that

$$8^N b\xi^{2-N} 2^{(N-1)(N-2)} \left[\frac{1}{16} \bar{\delta} 6^{-(N-1)} b_2 \right]^{2-N} \varrho_1^N \leq \frac{1}{64^N} \varrho_o^N.$$

This suggests the choice

$$\varrho_1 = \lambda \varrho_o \quad \text{with} \quad \lambda \stackrel{\text{def}}{=} \frac{(\bar{\delta}\xi)^{\frac{N-2}{N}}}{\gamma(N)b^{\frac{1}{N}}}.$$

Note that λ depends only on the data $\{N, C_o, C_1\}$, and not on ν . Hence, we can conclude that

$$Q_{8\varrho_1}^{\text{osc}}(b\xi^{2-N}\bar{\theta}_1) u \leq \omega_1,$$

which at this stage plays the role of (4.1).

Repeating the arguments of Sections 4.1–4.4 with the cylinder $Q_{8\varrho_1}(b\xi^{2-N}\bar{\theta}_1)$, we obtain that either

$$Q_{\frac{1}{8}\varrho_1}^{\text{osc}}(\theta_1) u \leq \left(1 - \frac{\bar{\eta}}{|\ln(\bar{\xi}\omega_1)|^{N-1}} \right) \omega_1, \quad \text{where} \quad \theta_1 = \left(\frac{\omega_1}{|\ln(\bar{\xi}\omega_1)|^{N-1}} \right)^{2-N}$$

or

$$\omega_1 \leq 8\epsilon.$$

Now, for every $n \in \mathbb{N}_0$ we may construct

$$\begin{aligned} \varrho_o &= 8\varrho, \quad \varrho_{n+1} = \lambda\varrho_n = \lambda^{n+1}\varrho_o = \lambda^{n+1}8\varrho, \\ \omega_o &= \omega, \quad \omega_{n+1} = \left(1 - \frac{\bar{\eta}}{|\ln(\bar{\xi}\omega_n)|^{N-1}} \right) \omega_n, \\ \bar{\theta}_n &= \left[\left(\frac{1}{8}\omega_n \right) \bar{\delta} \text{cap}_N \left(K_{\bar{c}\bar{c}_1^2(\frac{1}{8}\omega_n)^6}, K_3 \right) \right]^{2-N}, \quad \theta_n = \left(\frac{\omega_n}{|\ln(\bar{\xi}\omega_n)|^{N-1}} \right)^{2-N}, \\ Q'_n &= Q_{8\varrho_n}(b\xi^{2-N}\bar{\theta}_n), \quad Q_n = Q_{\frac{1}{8}\varrho_n}(\theta_n). \end{aligned}$$

By induction, if up to some $j \in \mathbb{N}$ we have

$$\omega_n > 8\epsilon, \quad \forall n \in \{0, 1, \dots, j-1\},$$

then, for all $n \in \{0, 1, \dots, j\}$, there holds

$$Q_n \subset Q'_n \subset Q_{n-1}, \quad \text{osc}_{Q_n} u \leq \text{osc}_{Q'_n} u \leq \omega_n.$$

On the other hand, we denote by j the first index to satisfy

$$\omega_j \leq 8\epsilon. \tag{4.20}$$

For $n \in \mathbb{N}_0$, consider the sequence

$$a_n = \exp(-c_* n^{\frac{1}{N}}), \quad \text{with} \quad c_* = \frac{\bar{\eta}}{2^{N-2} |\ln \bar{\xi}|^{N-1}}.$$

In Appendix B, we prove that

$$\begin{cases} a_{n+1} \geq a_n \left(1 - \frac{\bar{\eta}}{|\ln(\bar{\xi} a_n)|^{N-1}} \right), \\ a_o \geq \omega_o. \end{cases}$$

Hence, for any $n \in \mathbb{N}_0$, we have $a_n \geq \omega_n$.

Let us now take $r \in (0, \varrho)$; if, for some $n \in \{0, 1, \dots, j\}$, we have $\varrho_{n+1} \leq 8r < \varrho_n$, then

$$\operatorname{osc}_{Q_r(\theta_o)} u \leq \operatorname{osc}_{Q_n} u \leq \omega_n \leq \exp\left(-c_* n^{\frac{1}{N}}\right).$$

On the other hand, we have $\varrho_{n+1} = \lambda^{n+1} \varrho$. Therefore, $\varrho_{n+1} \leq 8r$ yields

$$n \geq \frac{1}{2|\ln \lambda|} \left| \ln \left(\frac{8r}{\varrho} \right) \right|,$$

and we conclude that

$$\begin{aligned} \operatorname{osc}_{Q_r(\theta_o)} u &\leq \exp\left(-\frac{c_*}{2^{\frac{1}{N}} |\ln \lambda|^{\frac{1}{N}}} \left| \ln \left(\frac{8r}{\varrho} \right) \right|^{\frac{1}{N}}\right) \\ &= \exp\left(-c \left| \ln \left(\frac{8r}{\varrho} \right) \right|^{\frac{1}{N}}\right), \quad \text{where } c \stackrel{\text{def}}{=} \frac{c_*}{2^{\frac{1}{N}} |\ln \lambda|^{\frac{1}{N}}}. \end{aligned}$$

It is not hard to trace the dependence of c on $\nu_* = \max\{1, \nu\}$ as

$$c = \frac{\gamma(N, C_o, C_1)}{\ln(2\nu_*)^{N-1}}. \quad (4.21)$$

Therefore, if we know ν_* is bounded by a larger number $\bar{\nu}$, then we may replace ν_* by $\bar{\nu}$ in the definition of c while retaining the same oscillation estimate.

On the other hand, if $r < \varrho_{j+1}$, where j is the first index for which (4.20) holds, we may use it and conclude that

$$\operatorname{osc}_{Q_r(\theta_o)} u \leq \operatorname{osc}_{Q_j} u \leq \omega_j \leq 8\epsilon.$$

This allows us to include the ϵ -term into the oscillation estimate and obtain that

$$\operatorname{osc}_{Q_r(\theta_o)} u \leq \exp\left(-c \left| \ln \left(\frac{8r}{\varrho} \right) \right|^{\frac{1}{N}}\right) + 8\epsilon.$$

Finally, we can let $\epsilon \rightarrow 0$ and conclude with the wanted modulus of continuity.

4.6. Modulus of continuity over compact sets. The gist of the previous arguments is that if there are parameters $\varrho > 0$ and $\omega_o \in (0, 1)$ such that the initial oscillation estimate (4.1) is verified, then the oscillation decay

$$\operatorname{osc}_{Q_r(\theta_o)} u \leq \exp\left(-c \left| \ln \left(\frac{8r}{\varrho} \right) \right|^{\frac{1}{N}}\right)$$

can be reached for all $r \in (0, \varrho)$. The initial oscillation estimate (4.1) can be rewritten as

$$\operatorname{osc}_{Q_{8\varrho}(L\theta_o)} u \leq \omega_o, \quad (4.22)$$

where $L > 1$ can be easily calculated in terms of the data $\{N, C_o, C_1\}$, making use of (4.10). Moreover, the condition $\omega_o \leq 1$ yields

$$1 < \theta_o = \left(\frac{\omega_o}{|\ln(\bar{\xi}\omega_o)|^{N-1}} \right)^{2-N},$$

and hence, under the assumption (4.22), we conclude that

$$\operatorname{osc}_{Q_r} u \leq \exp \left(-c \left| \ln \left(\frac{8r}{\varrho} \right) \right|^{\frac{1}{N}} \right). \quad (4.23)$$

We want to show how such a modulus of continuity is affected as we approach the parabolic boundary of E_T .

Let \mathcal{K} be a compact subset of E_T , and define

$$\mathbb{M} \stackrel{\text{def}}{=} \max \left\{ 1, \operatorname{osc}_{E_T} u \right\}, \quad 8R \stackrel{\text{def}}{=} p - \operatorname{dist}_{\text{par}}(\mathcal{K}, \partial_{\text{par}} E_T).$$

By the definition of p -parabolic distance, $\forall (x_o, t_o) \in \mathcal{K}$, we have

$$(x_o, t_o) + Q_{8R}(\mathbb{M}^{2-N}) \subset E_T.$$

Fix such a point $(x_o, t_o) \in \mathcal{K}$ and consider a new function on Q_{8R} defined by

$$v(x, t) \stackrel{\text{def}}{=} \frac{u(x - x_o, \mathbb{M}^{2-N}(t - t_o))}{\mathbb{M}}.$$

Then, apparently, we have

$$\operatorname{osc}_{Q_{8R}} v \leq 1.$$

Moreover, such v satisfies the same kind of Stefan problem (1.1), with structure conditions (1.3) but with a different jump constant ν/\mathbb{M} that defines $\beta(\cdot)$ in (1.2). The last oscillation estimate of v yields an analog of (4.22) if we let $\omega_o = 1$ and $\varrho = \sigma R$, where

$$\sigma = \frac{1}{L^{\frac{1}{N}} |\ln \bar{\xi}|^{\frac{(N-2)(N-1)}{N}}}$$

depends on $\{\nu, N, C_o, C_1\}$. Therefore, we may reproduce all previous arguments for v now. Moreover, due to $\mathbb{M} \geq 1$, the jump constant ν/\mathbb{M} that appears in all estimates regarding v can be replaced by the larger constant ν , cf. (4.21). As a result, we obtain an analog of oscillation decay (4.23) for v , with the constant c still determined by the data $\{\nu, N, C_o, C_1\}$. Reverting to u , we arrive at

$$\operatorname{osc}_{(x_o, t_o) + Q_r(\mathbb{M}^{2-N})} u \leq \mathbb{M} \exp \left(-c \left| \ln \left(\frac{8r}{\sigma R} \right) \right|^{\frac{1}{N}} \right),$$

for any $r \in (0, \sigma R)$.

Next, pick a second point $(x_1, t_1) \in \mathcal{K}$. Assume with no loss of generality that $t_1 \leq t_o$ and

$$r \stackrel{\text{def}}{=} |x_o - x_1| + \mathbb{M}^{\frac{N-2}{N}} (t_o - t_1)^{\frac{1}{N}} < \sigma R,$$

such that

$$(x_1, t_1) \in (x_o, t_o) + Q_r(\mathbb{M}^{2-N}).$$

The oscillation decay then yields

$$|u(x_o, t_o) - u(x_1, t_1)| \leq \mathbb{M} \exp \left(-c \left| \ln \left(\frac{|x_o - x_1| + \mathbb{M}^{\frac{N-2}{N}} (t_o - t_1)^{\frac{1}{N}}}{\frac{1}{8} \sigma R} \right) \right|^{\frac{1}{N}} \right).$$

This is the desired modulus of continuity over the compact set \mathcal{K} .

4.7. Hölder modulus of continuity. In this section, we briefly indicate how to modify the argument of the last subsections in order to obtain the Hölder continuity of weak solutions in the case $p > \max\{2, N\}$.

As before, the starting point is an oscillation estimate like (4.1). Indeed, suppose there are parameters $\varrho > 0$ and $\omega \in (0, 1)$ such that

$$Q_{8\varrho}(b\xi^{2-p}\bar{\theta}) \subset E_T, \quad \operatorname{osc}_{Q_{8\varrho}(b\xi^{2-p}\bar{\theta})} u \leq \omega, \quad (4.24)$$

where

$$\bar{\theta} \stackrel{\text{def}}{=} \left[\left(\frac{1}{8} \omega \right) \bar{\delta} b_1 \right]^{2-p}, \quad \alpha \stackrel{\text{def}}{=} c_1 \left(\frac{1}{8} \omega \right)^{\frac{N+2p}{p}}. \quad (4.25)$$

The various parameters, b , c_1 , $\bar{\delta}$, ξ , etc... retain the same token as in (4.1) and (4.2). However, for $\bar{\theta}$ in (4.25), in the place of the capacity, we employ its lower bound b_1 from (A.3), which depends only on $\{N, p\}$. For α , we change the power in accordance with (2.5).

Next, we introduce μ^\pm and consider, analogously, the two cases (4.3) and (4.4). Clearly, we define the interval I with the parameter N in the power replaced by p , whereas the two alternatives (4.5) and (4.6) remain the same upon taking these changes into consideration.

Section 4.1 can be reproduced along the same lines. Clearly, the parameter N that appears in the power of various quantities must be changed to p . As a result, an analog of the pointwise estimate (4.7) is reached in the cylinder (4.8). This will yield an oscillation estimate provided we choose $\tilde{\xi}$ and ξ properly, to guarantee

$$\frac{\delta}{2^p} \left(\frac{1}{8} \omega \right)^{2-p} \varrho^p + \gamma_o(\tilde{\xi}\omega)^{2-p} \frac{1}{4^p} \varrho^p \geq 2b\xi^{2-p}\bar{\theta}\varrho^p, \quad (4.26)$$

$$-b\xi^{2-p}\bar{\theta}\varrho^p + \frac{\delta}{2^p} \left(\frac{1}{8} \omega \right)^{2-p} \varrho^p < -\theta \left(\frac{1}{8} \varrho \right)^p, \quad (4.27)$$

where

$$\theta = \omega^{2-p}. \quad (4.28)$$

Discarding the term with δ in (4.26) and considering $\bar{\theta}$ defined in (4.25), the choice of $\tilde{\xi}$ now becomes

$$\tilde{\xi} \leq \left(\frac{\gamma_o}{b} \right)^{\frac{1}{p-2}} 2^{\frac{5(1-p)}{p-2}} \xi \bar{\delta} b_1,$$

whereas (4.27) is satisfied provided ξ is small enough. Therefore, we arrive at an analog of the oscillation estimate (4.13), with θ defined in (4.28) and with η replaced by $\frac{1}{2}\tilde{\xi}$, which depends only on the data $\{N, p, C_o, C_1\}$ and is independent of ω and ν .

Section 4.2 can also be reproduced along the same lines. However, in the application of Proposition 3.1 to v , we obtain instead that

$$v \geq \frac{1}{8} \omega \xi b_1$$

in $K_{\frac{1}{2}\varrho} \times (-b\xi^{2-p}\bar{\theta}\varrho^p, 0]$, because of the capacity estimate in (A.3). Consequently, we obtain an analog of the oscillation estimate (4.15), with θ defined in (4.28) and with η replaced by $\frac{1}{8}\xi b_1$, which depends only on the data $\{N, p, C_o, C_1\}$, and is independent of ω and ν .

Necessary changes can be performed similarly for Sections 4.3 and 4.4. Therefore, under the assumption that (4.24) holds true for $\varrho > 0$ and $\omega \in (0, 1)$, we eventually obtain

$$\operatorname{osc}_{Q_{\frac{1}{8}\varrho}(\theta)} u \leq (1 - \eta)\omega, \quad \text{with } \theta = \omega^{2-p}.$$

Note that $\eta = \min\{\frac{1}{2}\tilde{\xi}, \frac{1}{8}\xi b_1\}$ depends only on $\{N, p, C_o, C_1\}$, but neither on ω nor on ν . Setting up an iteration scheme and deriving an Hölder modulus of continuity now become standard.

Finally, it is apparent that when $2 < p < N$, $N \geq 3$, Proposition 3.1 yields a power-like dependence on α of the pointwise estimate. This will affect the recurrence of ω_n and, consequently, the modulus of continuity. Once this is considered, the previous computations can also be appropriately adapted for this case. All the moduli of Table 1 are now justified.

APPENDIX A. SOME PROPERTIES OF p -CAPACITY

For $p \geq 1$, the notion of p -capacity is defined by

$$\operatorname{cap}_p(F, \Omega) \stackrel{\text{def}}{=} \inf_{u \in W} \int_{\Omega} |Du|^p dx,$$

where F is a compact subset of the open set Ω in \mathbb{R}^N and

$$W \stackrel{\text{def}}{=} \{u \in C_o^\infty(\Omega) : u \geq 1 \text{ on } F\}.$$

After proper approximations, the above minimization could take place over $W_o^{1,p}(\Omega)$ instead of $C_o^\infty(\Omega)$, and the p -capacity of an arbitrary subset of Ω can be formulated based on that of compact subsets. The reader is referred to [15, Chapter 2].

Immediate from the definition is the scaling property

$$\frac{\operatorname{cap}_p(K_{\varepsilon\varrho}, K_\varrho)}{\varrho^{N-p}} = \operatorname{cap}_p(K_\varepsilon, K_1), \quad \text{for } \varepsilon \in (0, 1). \quad (\text{A.1})$$

Also immediate is that

$$\operatorname{cap}_p(F, \Omega) \geq \operatorname{cap}_p(F, \Omega_o),$$

if Ω_o is an open set satisfying $F \subset \Omega \subset \Omega_o$. This, in particular, implies that

$$\operatorname{cap}_p(K_\varepsilon, K_1) \geq \operatorname{cap}_p(K_\varepsilon, K_3).$$

With a little more effort, one also shows that the reverse estimate holds, apart from a multiplicative constant depending on N ; see [15, § 2.16].

Capacities of balls can be estimated explicitly. Namely,

$$\operatorname{cap}_p(B_r, B_R) = \begin{cases} \omega_{N-1} \left(\frac{|N-p|}{p-1}\right)^{p-1} \left| R^{\frac{p-N}{p-1}} - r^{\frac{p-N}{p-1}} \right|^{1-p} & \text{if } p \neq N, \\ \omega_{N-1} \left| \ln \frac{r}{R} \right|^{1-N} & \text{if } p = N, \end{cases} \quad (\text{A.2})$$

where ω_{N-1} denotes the surface measure of the boundary of the unit ball in \mathbb{R}^N . See [15, § 2.11]. Based on this, one easily estimates

$$\text{cap}_p(K_\varepsilon, K_3) \leq \gamma(N, p), \quad (\text{A.3})$$

for any $\varepsilon \in (0, 1)$.

Another fact that can be derived from (A.2) with little effort is

$$\text{cap}_p(K_\varepsilon, K_3) \geq \begin{cases} b_1 & \text{if } p > N, \\ b_2 |\ln \varepsilon|^{-(N-1)} & \text{if } p = N, \\ b_3 \varepsilon^{N-p} & \text{if } p < N, \end{cases} \quad (\text{A.4})$$

for positive constants b_1 , b_2 and b_3 depending only on $\{N, p\}$, and the lower bound is actually an equality when $p = N$.

APPENDIX B. ESTIMATE OF A PARAMETER

Consider the decreasing sequence

$$a_n = \exp(-c_* n^{\frac{1}{N}}), \quad n \in \mathbb{N}_0.$$

Our goal is to select the positive parameter c_* in such a way that, for any $n \in \mathbb{N}_0$,

$$\begin{cases} a_{n+1} \geq a_n \left(1 - \frac{\bar{\eta}}{|\ln(\bar{\xi} a_n)|^{N-1}}\right), \\ a_o \geq \omega, \end{cases} \quad (\text{B.1})$$

where $\bar{\eta}$, $\bar{\xi}$, and ω are the same quantities as in (4.17).

Since $a_o = 1$, and in the derivation of (4.17), we assumed $\omega \leq 1$, the second condition of (B.1) is automatically satisfied.

Moreover, we have

$$\begin{aligned} 1 - \frac{\bar{\eta}}{|\ln(\bar{\xi} a_n)|^{N-1}} &= 1 - \frac{\bar{\eta}}{[|\ln a_n| + |\ln \bar{\xi}|]^{N-1}} \\ &= 1 - \frac{\bar{\eta}}{[c_* n^{\frac{1}{N}} + |\ln \bar{\xi}|]^{N-1}} \leq \exp\left(-\frac{\bar{\eta}}{[c_* n^{\frac{1}{N}} + |\ln \bar{\xi}|]^{N-1}}\right), \end{aligned}$$

where we have taken into account that one has $1 - x \leq e^{-x}$, $\forall x \in \mathbb{R}$.

On the other hand, since

$$\frac{a_{n+1}}{a_n} = \exp\left(-c_*[(n+1)^{\frac{1}{N}} - n^{\frac{1}{N}}]\right),$$

it suffices to choose c_* such that

$$c_*[(n+1)^{\frac{1}{N}} - n^{\frac{1}{N}}] \leq \frac{\bar{\eta}}{[c_* n^{\frac{1}{N}} + |\ln \bar{\xi}|]^{N-1}},$$

that is

$$c_*[(n+1)^{\frac{1}{N}} - n^{\frac{1}{N}}][c_* n^{\frac{1}{N}} + |\ln \bar{\xi}|]^{N-1} \leq \bar{\eta}.$$

Since, for any $a, b > 0$, we have $(a + b)^{N-1} \leq 2^{N-2}(a^{N-1} + b^{N-1})$, we require

$$\begin{aligned} 2^{N-2}c_*[(n+1)^{\frac{1}{N}} - n^{\frac{1}{N}}][c_*^{N-1}n^{\frac{N-1}{N}} + |\ln \bar{\xi}|^{N-1}] &\leq \bar{\eta} \\ c_*^N[(n^N + n^{N-1})^{\frac{1}{N}} - n] + c_*[(n+1)^{\frac{1}{N}} - n^{\frac{1}{N}}]|\ln \bar{\xi}|^{N-1} &\leq \frac{\bar{\eta}}{2^{N-2}}. \end{aligned}$$

If $n = 0$, we require

$$c_* \leq \frac{\bar{\eta}}{2^{N-2}|\ln \bar{\xi}|^{N-1}}.$$

If $n \geq 1$, recalling that, for any $m \in (0, 1)$ and for any $x > -1$, we have $(1 + x)^m \leq 1 + mx$, yields

$$\begin{aligned} (n^N + n^{N-1})^{\frac{1}{N}} - n &= n\left(1 + \frac{1}{n}\right)^{\frac{1}{N}} - n \leq n\left(1 + \frac{1}{N}\frac{1}{n}\right) - n = \frac{1}{N}, \\ (n+1)^{\frac{1}{N}} - n^{\frac{1}{N}} &= n^{\frac{1}{N}}\left(1 + \frac{1}{n}\right)^{\frac{1}{N}} - n^{\frac{1}{N}} \leq n^{\frac{1}{N}} + \frac{1}{N}\frac{1}{n^{\frac{N-1}{N}}} - n^{\frac{1}{N}} \leq \frac{1}{N}. \end{aligned}$$

Hence, we require

$$\begin{aligned} c_*^N + c_*|\ln \bar{\xi}|^{N-1} &\leq \frac{N\bar{\eta}}{2^{N-2}}, \\ 2c_*|\ln \bar{\xi}|^{N-1} &\leq \frac{N\bar{\eta}}{2^{N-2}}, \\ c_* &\leq \frac{N\bar{\eta}}{2^{N-1}|\ln \bar{\xi}|^{N-1}}. \end{aligned}$$

Therefore, in order to have c_* that satisfies (B.1) for any $n \geq 0$, it suffices to take

$$c_* \stackrel{\text{def}}{=} \min \left\{ \frac{N\bar{\eta}}{2^{N-1}|\ln \bar{\xi}|^{N-1}}, \frac{\bar{\eta}}{2^{N-2}|\ln \bar{\xi}|^{N-1}} \right\} = \frac{\bar{\eta}}{2^{N-2}|\ln \bar{\xi}|^{N-1}}.$$

Acknowledgments. UG is partially supported by the grant 202244A7YL ‘‘Gradient flows and non-smooth geometric structures with applications to optimization and machine learning.’’ NL is supported by the FWF-project P36272-N ‘‘On the Stefan type problems.’’ JMU is partially supported by KAUST and CMUC (funded by the Portuguese Government through FCT/MCTES, DOI 10.54499/UIDB/00324/2020). Part of this paper was written while the authors participated in the Workshop on Nonlinear Parabolic PDEs, held at the Institut Mittag-Leffler in May 2024.

REFERENCES

- [1] P. Baroni, T. Kuusi and J.M. Urbano, *A quantitative modulus of continuity for the two-phase Stefan problem*, Arch. Ration. Mech. Anal. **214** (2014), no. 2, 545–573.
- [2] V. Bögelein, F. Duzaar, R. Giova and A. Passarelli di Napoli, *Higher regularity in congested traffic dynamics*, Math. Ann. **385** (2023), no. 3–4, 1823–1878.
- [3] V. Bögelein, F. Duzaar, R. Giova and A. Passarelli di Napoli, *Gradient regularity for a class of widely degenerate parabolic systems*, SIAM J. Math. Anal. **56** (2024), no. 4, 5017–5078.
- [4] L.A. Caffarelli and L.C. Evans, *Continuity of the temperature in the two-phase Stefan problem*, Arch. Ration. Mech. Anal. **81** (1983), no. 3, 199–220.
- [5] L.A. Caffarelli and A. Friedman, *Continuity of the temperature in the Stefan problem*, Indiana Univ. Math. J. **28** (1979), no. 1, 53–70.

- [6] E. DiBenedetto, *Continuity of weak solutions to certain singular parabolic equations*, Ann. Mat. Pura Appl.(4) **130** (1982), 131–176.
- [7] E. DiBenedetto, *A boundary modulus of continuity for a class of singular parabolic equations*, J. Differential Equations **63** (1986), no. 3, 418–447.
- [8] E. DiBenedetto, “Degenerate Parabolic Equations”, Universitext, Springer-Verlag, New York, 1993.
- [9] E. DiBenedetto and A. Friedman, *Regularity of solutions of nonlinear degenerate parabolic systems*, J. Reine Angew. Math. **349** (1984), 83–128.
- [10] E. DiBenedetto, U. Gianazza and V. Vespi, “Harnack’s Inequality for Degenerate and Singular Parabolic Equations”, Springer Monographs in Mathematics, Springer-Verlag, New York, 2012.
- [11] E. DiBenedetto and V. Vespi, *On the singular equation $\beta(u)_t = \Delta u$* , Arch. Ration. Mech. Anal. **132** (1995), no. 3, 247–309.
- [12] U. Gianazza and N. Liao, *A boundary estimate for degenerate parabolic diffusion equations*, Potential Anal. **53** (2020), no. 3, 977–995.
- [13] U. Gianazza and N. Liao, *Continuity of the temperature in a multi-phase transition problem*, Math. Ann. **384** (2022), no. 1-2, 211–245.
- [14] U. Gianazza, M. Surnachev and V. Vespi, *On a new proof of Hölder continuity of solutions of p -Laplace type parabolic equations*, Adv. Calc. Var. **3** (2010), no. 3, 263–278.
- [15] J. Heinonen, T. Kilpeläinen and O. Martio, “Nonlinear Potential Theory of Degenerate Elliptic Equations”, 2nd ed., Dover, Mineola, NY, 2006.
- [16] T. Kuusi, *Harnack estimates for weak supersolutions to nonlinear degenerate parabolic equations*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **7** (2008), no. 4, 673–716.
- [17] N. Liao, *On the logarithmic type boundary modulus of continuity for the Stefan problem: to the memory of Emmanuele DiBenedetto*, Adv. Math. **408** (2022), Paper No. 108613, 53 pp.
- [18] N. Liao, *An improved modulus of continuity for the two-phase Stefan problem*, Trans. Amer. Math. Soc. (2024), 19 pp.
- [19] N. Liao, *Hölder estimates for the Stefan problem*, preprint (2023), 1–20.
- [20] O.A. Oleĭnik, *A method of solution of the general Stefan problem*, Soviet Math. Dokl. **1**((1960), 1350–1354.
- [21] P.E. Sacks, *Continuity of solutions of a singular parabolic equation*, Nonlinear Anal. **7** (1983), no. 4, 387–409.
- [22] J.M. Urbano, *Continuous solutions for a degenerate free boundary problem*, Ann. Mat. Pura Appl.(4) **178** (2000), 195–224.
- [23] J.M. Urbano, “The method of intrinsic scaling”, Lecture Notes in Math., 1930, Springer-Verlag, Berlin, 2008.
- [24] A. Visintin, *Introduction to Stefan-type problems*, “Handbook of differential equations: evolutionary equations. Vol. IV”, 377–484, Handb. Differ. Equ., Elsevier/North-Holland, Amsterdam, 2008.
- [25] W.P. Ziemer, *Interior and boundary continuity of weak solutions of degenerate parabolic equations*, Trans. Amer. Math. Soc. **271** (1982), no. 2, 733–748.

DIPARTIMENTO DI MATEMATICA “F. CASORATI”, UNIVERSITÀ DI PAVIA, VIA FERRATA 5, 27100 PAVIA, ITALY

Email address: `ugogia04@unipv.it`

FACHBEREICH MATHEMATIK, UNIVERSITÄT SALZBURG, HELLBRUNNER STR. 34, 5020 SALZBURG, AUSTRIA

Email address: `naian.liao@plus.ac.at`

APPLIED MATHEMATICS AND COMPUTATIONAL SCIENCES PROGRAM (AMCS), COMPUTER, ELECTRICAL AND MATHEMATICAL SCIENCES AND ENGINEERING DIVISION (CEMSE), KING ABDULLAH UNIVERSITY OF SCIENCE AND TECHNOLOGY (KAUST), THUWAL, 23955 -6900, KINGDOM OF SAUDI ARABIA AND CMUC, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF COIMBRA, 3000-143 COIMBRA, PORTUGAL

Email address: `miguel.urbano@kaust.edu.sa`